PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47, A61K 38/17 — A2

(11) International Publication Number: WO 98/55614

(43) International Publication Date: 10 December 1998 (10.12.98)

(21) International Application Number: PCT/US98/11210

(22) International Filing Date: 1 June 1998 (01.06.98)

(30) Priority Data:

08/868,899	4 June 1997 (04.06.97)	US
08/868,898	4 June 1997 (04.06.97)	US
08/869,192	4 June 1997 (04.06.97)	US
08/869,191	4 June 1997 (04.06.97)	US
08/869,193	4 June 1997 (04.06.97)	US
08/868,697	4 June 1997 (04.06.97)	US
08/868,698	4 June 1997 (04.06.97)	US
08/868,900	4 June 1997 (04.06.97)	US
08/868,696	4 June 1997 (04.06.97)	US
08/869,194	4 June 1997 (04.06.97)	US
09/087,255	29 May 1998 (29.05.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Street, Newton, MA 02160 (US). McCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US). HOWES, Steven, H.; Apartment 2, 44 Chester Street, Somerville, MA 02144 (US). FECHTEL, Kim; 46 Marion Road, Arlington, MA 02174 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of the following applications:

(1) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/868,899), filed June 4, 1997;

(2) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/868,898), filed June 4, 1997;

(3) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/869,192), filed June 4, 1997;

(4) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/869,191), filed June 4, 1997;

(5) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/869,193), filed June 4, 1997;

(6) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/868,697), filed June 4, 1997;

(7) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/868,698), filed June 4, 1997;

(8) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/868,900), filed June 4, 1997;

(9) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/868,696), filed June 4, 1997;

(10) Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/869,194), filed June 4, 1997;

all of which are incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 266 to nucleotide 1651;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 521 to nucleotide 1651;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 335 to nucleotide 634;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone as294_3 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone as294_3 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone as294_3 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone as294_3 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 266 to nucleotide 1651; the nucleotide sequence of SEQ ID NO:1 from nucleotide 521 to nucleotide 1651; the nucleotide sequence of SEQ ID NO:1 from nucleotide 335 to nucleotide 634; the nucleotide sequence of the full-length protein coding sequence of clone as294_3 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone as294_3 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone as294_3 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 123. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2

having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:2, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 226 to amino acid 235 of SEQ ID NO:2.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 123;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising eight consecutive amino acids of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone as294_3 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 123. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:2, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2having biological activity, the fragment comprising the amino acid sequence from amino acid 226 to amino acid 235 of SEQ ID NO:2.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 262 to nucleotide 3096;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1118 to nucleotide 1527;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone aw92_1 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone aw92_1 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone aw92_1 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone aw92_1 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 262 to nucleotide 3096; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1118 to nucleotide 1527; the nucleotide sequence of the full-length protein coding sequence of clone aw92_1 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone aw92_1 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone aw92_1 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 287 to amino acid 422. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably

twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:4, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 467 to amino acid 476 of SEQ ID NO:4.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 287 to amino acid 422;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising eight consecutive amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone aw92_1 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 287 to amino acid 422. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:4, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4having biological activity, the fragment comprising the amino acid sequence from amino acid 467 to amino acid 476 of SEQ ID NO:4.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 612 to nucleotide 806;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 744 to nucleotide 806;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 794;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bd316_2 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bd316_2 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bd316_2 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bd316_2 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 612 to nucleotide 806; the nucleotide sequence of SEQ ID NO:5 from nucleotide 744 to nucleotide 806; the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 794; the nucleotide sequence of the full-length protein coding sequence of clone bd316_2 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone bd316_2 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bd316_2 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 61.

In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:6, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 27 to amino acid 36 of SEQ ID NO:6.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 61;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising eight consecutive amino acids of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone bd316_2 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 61. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:6, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6having biological activity, the fragment comprising the amino acid sequence from amino acid 27 to amino acid 36 of SEQ ID NO:6.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 7 to nucleotide 300;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 363;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk130_4 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk130_4 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bk130_4 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bk130_4 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 7 to nucleotide 300; the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 363; the nucleotide sequence of the full-length protein coding sequence of clone bk130_4 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone bk130_4 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bk130_4 deposited under accession number ATCC 98444. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most

preferably thirty) consecutive amino acids of SEQ ID NO:8, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 44 to amino acid 53 of SEQ ID NO:8.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising eight consecutive amino acids of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone bk130_4 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:8, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 44 to amino acid 53 of SEQ ID NO:8.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 52 to nucleotide 1863;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1219 to nucleotide 1863;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1099 to nucleotide 1743;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bv131_5 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bv131_5 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bv131_5 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bv131_5 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:10;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 52 to nucleotide 1863; the nucleotide sequence of SEQ ID NO:9 from nucleotide 1219 to nucleotide 1863; the nucleotide sequence of SEQ ID NO:9 from nucleotide 1099 to nucleotide 1743; the nucleotide sequence of the full-length protein coding sequence of clone bv131_5 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone bv131_5 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bv131_5 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10 from amino acid 430 to amino acid 564. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10

having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:10, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 297 to amino acid 306 of SEQ ID NO:10.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 430 to amino acid 564;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising eight consecutive amino acids of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone bv131_5 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:10 from amino acid 430 to amino acid 564. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:10, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10having biological activity, the fragment comprising the amino acid sequence from amino acid 297 to amino acid 306 of SEQ ID NO:10.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 67 to nucleotide 690;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 1 to nucleotide 576;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bv227_1 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bv227_1 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bv227_1 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bv227_1 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:12;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 67 to nucleotide 690; the nucleotide sequence of SEQ ID NO:11 from nucleotide 1 to nucleotide 576; the nucleotide sequence of the full-length protein coding sequence of clone bv227_1 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone bv227_1 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bv227_1 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 170. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment preferably comprising eight (more preferably

twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:12, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 99 to amino acid 108 of SEQ ID NO:12.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 170;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising eight consecutive amino acids of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone bv227_1 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 170. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:12, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12having biological activity, the fragment comprising the amino acid sequence from amino acid 99 to amino acid 108 of SEQ ID NO:12.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 657 to nucleotide 1469;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 678 to nucleotide 1103;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cd265_11 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cd265_11 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cd265_11 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cd265_11 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:14;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 657 to nucleotide 1469; the nucleotide sequence of SEQ ID NO:13 from nucleotide 678 to nucleotide 1103; the nucleotide sequence of the full-length protein coding sequence of clone cd265_11 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone cd265_11 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone cd265_11 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 8 to amino acid 149. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment preferably comprising eight (more preferably

twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:14, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 130 to amino acid 139 of SEQ ID NO:14.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 8 to amino acid 149;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising eight consecutive amino acids of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone cd265_11 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence of SEQ ID NO:14 from amino acid 8 to amino acid 149. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:14, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14having biological activity, the fragment comprising the amino acid sequence from amino acid 130 to amino acid 139 of SEQ ID NO:14.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 261 to nucleotide 896;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 330 to nucleotide 896;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 515;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ej265_4 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ej265_4 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ej265_4 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ej265_4 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:16;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 261 to nucleotide 896; the nucleotide sequence of SEQ ID NO:15 from nucleotide 330 to nucleotide 896; the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 515; the nucleotide sequence of the full-length protein coding sequence of clone ej265_4 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone ej265_4 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ej265_4 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 85. In further preferred

embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:16, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 101 to amino acid 110 of SEQ ID NO:16.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 85;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising eight consecutive amino acids of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone ej265_4 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 85. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:16, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16having biological activity, the fragment comprising the amino acid sequence from amino acid 101 to amino acid 110 of SEQ ID NO:16.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 946 to nucleotide 2232;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 1336 to nucleotide 1853;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ey29_8 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ey29_8 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ey29_8 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ey29_8 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:18;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 946 to nucleotide 2232; the nucleotide sequence of SEQ ID NO:17 from nucleotide 1336 to nucleotide 1853; the nucleotide sequence of the full-length protein coding sequence of clone ey29_8 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone ey29_8 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ey29_8 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18 from amino acid 138 to amino acid 302. In further preferred embodiments, the present invention provides a polynucleotide encoding a

protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:18, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 209 to amino acid 218 of SEQ ID NO:18.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 138 to amino acid 302;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising eight consecutive amino acids of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone ey29_8 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18 or the amino acid sequence of SEQ ID NO:18 from amino acid 138 to amino acid 302. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:18, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18having biological activity, the fragment comprising the amino acid sequence from amino acid 209 to amino acid 218 of SEQ ID NO:18.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2588 to nucleotide 3439;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 3005 to nucleotide 3502;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gm114_10 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gm114_10 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gm114_10 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gm114_10 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:20;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 2588 to nucleotide 3439; the nucleotide sequence of SEQ ID NO:19 from nucleotide 3005 to nucleotide 3502; the nucleotide sequence of the full-length protein coding sequence of clone gm114_10 deposited under accession number ATCC 98444; or the nucleotide sequence of a mature protein coding sequence of clone gm114_10 deposited under accession number ATCC 98444. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone gm114_10 deposited under accession number ATCC 98444. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid 145 to amino acid 284. In further preferred embodiments, the present invention provides a

polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:20, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 137 to amino acid 146 of SEQ ID NO:20.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 145 to amino acid 284;

(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising eight consecutive amino acids of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone gm114_10 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence of SEQ ID NO:20 from amino acid 145 to amino acid 284. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:20, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20having biological activity, the fragment comprising the amino acid sequence from amino acid 137 to amino acid 146 of SEQ ID NO:20.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell

in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "as294_3"

A polynucleotide of the present invention has been identified as clone "as294_3". as294_3 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. as294_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "as294_3 protein").

The nucleotide sequence of as294_3 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the as294_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 73 to 85 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 86, or are a transmembrane domain. Amino acids 102 to 114 are also a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 115, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone as294_3 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for as294_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. as294_3 demonstrated at least some similarity with sequences identified as AA206777 (zq80d04.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone 647911 3'), AA206905 (zq80d04.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone 647911 5'), AA280222 (zt04c05.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE 712136 5'), H19869 (yn57a08.s1 Homo sapiens cDNA clone 172502 3'), H24249 (ym50h12.r1 Homo sapiens cDNA clone 52050 5'), N44936 (yy34f11.r1 Homo sapiens cDNA clone 273165 5'), R15379 (yf90f03.r1 Homo sapiens cDNA clone 29694 5'), R43727 (yg20c11.s1 Homo sapiens cDNA clone 32810 3'), R88673 (ym93f09.r1 Homo sapiens cDNA clone 166505 5'), T21648 (Human gene signature HUMGS03085), T80165 (5p IMAGE clone), and Z99260 (GenPept S. pombe hypothetical

protein). The predicted amino acid sequence disclosed herein for as294_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted as294_3 protein demonstrated at least some similarity to sequences identified as X73434 (KAP5.4 keratin protein [Ovis aries]) and Z99260 (hypothetical protein [Schizosaccharomyces pombe]). Based upon sequence similarity, as294_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts three potential transmembrane domains within the as294_3 protein sequence, centered around amino acids 105, 228, and 307 of SEQ ID NO:2, respectively.

Clone "aw92_1"

A polynucleotide of the present invention has been identified as clone "aw92_1". aw92_1 was isolated from a cDNA library of human adult ovary (comprising untreated tissue and tissue treated with retinoic acid and activin), using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. aw92_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "aw92_1 protein").

The nucleotide sequence of aw92_1 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the aw92_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone aw92_1 should be approximately 2950 bp.

The nucleotide sequence disclosed herein for aw92_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. aw92_1 demonstrated at least some similarity with sequences identified as AF021936 (Rattus norvegicus myotonic dystrophy kinase-related Cdc42-binding kinase MRCK-beta (MRCK-beta) mRNA, complete CDs, GP2736153), T23529 (seq3368 Homo sapiens cDNA clone Hy18-Charon40-cDNA-247 3'), U59305 (Human ser-thr protein kinase PK428 mRNA, complete cds), W16524 (zb15h09.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 302177 5' similar to PIR A42101 A42101 protein kinase homolog - human; contains element MER22 repetitive element), and

X69292 (H.sapiens mRNA for smooth muscle myosin). The predicted amino acid sequence disclosed herein for aw92_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted aw92_1 protein demonstrated at least some similarity to sequences identified as L03534 (ENHMHCAX_1 myosin heavy chain [Entamoeba histolytica]), R41000 (Human brain cDNA clone C28 protein kinase), U59305 (ser-thr protein kinase PK428 [Homo sapiens]), W02258 (Nucleolar/endosomal auto-antigen p162), and X03740 (myosin heavy chain (876 AA) [Homo sapiens]). Based upon sequence similarity, aw92_1 proteins and each similar protein or peptide may share at least some activity.

Clone "bd316_2"

A polynucleotide of the present invention has been identified as clone "bd316_2". bd316_2 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bd316_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bd316_2 protein").

The nucleotide sequence of bd316_2 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bd316_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 32 to 44 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 45, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bd316_2 should be approximately 1200 bp.

The nucleotide sequence disclosed herein for bd316_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bd316_2 demonstrated at least some similarity with sequences identified as AA234339 (zr72d12.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 668951 3'), L05367 (Human oligodendrocyte myelin glycoprotein (OMG) exons 1-2; neurofibromatosis 1 (NF1) exons 28-49; ecotropic viral integration site 2B (EVI2B) exons 1-2; ecotropic viral integration site 2A (EVI2A) exons 1-2; adenylate kinase (AK3) exons

1-2), N30778 (yw74h08.s1 Homo sapiens cDNA clone 258015 3' similar to gblM73048|HUMU3AAAA Human U3 small nuclear RNA (rRNA);contains MER12.t1 MER12 repetitive element), U52195 (Human desmoglein 3 gene, promoter region), U60822 (Human dystrophin (DMD) gene, exons 7, 8 and 9, and partial cds), X85184 (R.norvegicus mRNA for ras-related GTPase, ragB), and X90530 (H.sapiens mRNA for ragB protein). Based upon sequence similarity, bd316_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the bd316_2 protein sequence centered around amino acid 35 of SEQ ID NO:6.

Clone "bk130_4"

A polynucleotide of the present invention has been identified as clone "bk130_4". bk130_4 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bk130_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bk130_4 protein").

The nucleotide sequence of bk130_4 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bk130_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bk130_4 should be approximately 550 bp.

The nucleotide sequence disclosed herein for bk130_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bk130_4 demonstrated at least some similarity with sequences identified as AA009736 (ze82e04.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 365502 3'), AA112971 (zn59b09.r1 Stratagene muscle 937209 Homo sapiens cDNA clone 562457 5'), AA196543 (zq08e12.s1 Stratagene muscle 937209 Homo sapiens cDNA clone 629134 3'), AA196677 (zq08e10.r1 Stratagene muscle 937209 Homo sapiens cDNA clone 629130 5'), AA232667 (zr74e10.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 669162 3'), H26737 (yl14f12.r1 Homo sapiens cDNA clone 158255 5'), H44642

(yp20a08.r1 Homo sapiens cDNA clone 187958 5'), and W72771 (zd77c12.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 346678 5'). The predicted amino acid sequence disclosed herein for bk130_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bk130_4 protein demonstrated at least some similarity to sequences identified as L11647 (glycogen branching enzyme [Streptomyces aureofaciens]), L23651(homology with C. elegans cuticle collagen; putative [Caenorhabditis elegans]), W03740 (rchd528 gene product), and Z29095 (R10E11.1 [Caenorhabditis elegans]). Based upon sequence similarity, bk130_4 proteins and each similar protein or peptide may share at least some activity.

Clone "bv131_5"

A polynucleotide of the present invention has been identified as clone "bv131_5". bv131_5 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bv131_5 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bv131_5 protein").

The nucleotide sequence of bv131_5 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bv131_5 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10. Amino acids 377 to 389 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 390, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bv131_5 should be approximately 2900 bp.

The nucleotide sequence disclosed herein for bv131_5 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bv131_5 demonstrated at least some similarity with sequences identified as AA233510 (zr29h03.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 664853 5' similar to TR:G1151007 G1151007 ATP DEPENDENT PERMEASE), H24176 (ym55e05.r1 Homo sapiens cDNA clone 52176 5'), R13832 (yf65a02.r1 Homo sapiens cDNA clone 26986 5' similar to SP:ADP1_YEAST P25371

PROBABLE ATP-DEPENDENT PERMEASE), R16423 (yf40d03.r1 Homo sapiens cDNA clone 129317 5'), T00880 (Human cisplatin resistance gene cDNA62), T12316 (Replicable and transcriptionally active plasmid), T78871 (yd83b08.s1 Homo sapiens cDNA clone 114807 3'), U66681 (Human clone EST157481 ATP-binding cassette transporter mRNA sequence), and V00710 (Human mitochondrial genes for several tRNAs (Phe, Val, Leu) and 12S and 16S ribosomal RNAs). The predicted amino acid sequence disclosed herein for bv131_5 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bv131_5 protein demonstrated at least some similarity to sequences identified as U34919 (white homolog [Homo sapiens]), Z48745 (murine ABC8), and Z49821 (putative ABC transporter [Saccharomyces cerevisiae]). Based upon sequence similarity, bv131_5 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts five additional potential transmembrane domains within the bv131_5 protein sequence, centered around amino acids 354, 439, 463, 494 and 588 of SEQ ID NO:10, respectively.

Clone "bv227_1"

A polynucleotide of the present invention has been identified as clone "bv227_1". bv227_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bv227_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bv227_1 protein").

The nucleotide sequence of bv227_1 as presently determined is reported in SEQ ID NO:11. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bv227_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12. Amino acids 45 to 57 of SEQ ID NO:12 are a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 58, or are a transmembrane domain. Another potential bv227_1 reading frame and predicted amino acid sequence is encoded by basepairs 921 to 2294 of SEQ ID NO:11 and is reported in SEQ ID NO:31. A frameshift in the nucleotide sequence of SEQ ID NO:11 between about nucleotide 664 to about nucleotide 690 could extend the

reading frame of SEQ ID NO:31 to form a reading frame extending from position 666 to 2294 of SEQ ID NO:11 and encoding the amino acid sequence reported in SEQ ID NO:32.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bv227_1 should be approximately 3300 bp.

The nucleotide sequence disclosed herein for bv227_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bv227_1 demonstrated at least some similarity with sequences identified as AA368932 (EST80282 Placenta I Homo sapiens cDNA similar to similar to beta-1-glycoprotein PSGGA, pregnancy-specific), D60272 (Human fetal brain cDNA 3'-end GEN-095A07), M58526 (Human alpha-5 collagen type IV (COL4A5) mRNA, 3' end), Q64556 (Human collagen (Type V) coding sequence), R74388 (yi57f11.s1 Homo sapiens cDNA clone 143373 3'), and T67066 (Human alpha3(IX) collagen cDNA). The predicted amino acid sequences disclosed herein for bv227_1 were searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bv227_1 proteins of SEQ ID NO:31 and SEQ ID NO:32 demonstrated at least some similarity to sequences identified as S57132 (type XVI collagen alpha 1 chain, alpha 1 (XVI) [human, placenta, Peptide Partial, 1186 aa] [Homo sapiens]) and W07539 (Collagen like protein (CLP)). Based upon sequence similarity, bv227_1 proteins and each similar protein or peptide may share at least some activity.

Clone "cd265_11"

A polynucleotide of the present invention has been identified as clone "cd265_11". cd265_11 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cd265_11 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cd265_11 protein").

The nucleotide sequence of cd265_11 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cd265_11 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cd265_11 should be approximately 1600 bp.

The nucleotide sequence disclosed herein for cd265_11 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cd265_11 demonstrated at least some similarity with sequences identified as AA125395 (mp77f05.r1 Soares 2NbMT Mus musculus cDNA clone 575265 5'), AA131340 (zo08h01.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 567121 3'), AA244194 (nc06b11.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone 1462), AA339557 (EST44738 Fetal brain I Homo sapiens cDNA 5' end), AA569649 (nf24a11.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:914684), and T26052 (Human gene signature HUMGS08288). Based upon sequence similarity, cd265_11 proteins and each similar protein or peptide may share at least some activity.

Clone "ej265_4"

A polynucleotide of the present invention has been identified as clone "ej265_4". ej265_4 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ej265_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ej265_4 protein").

The nucleotide sequence of ej265_4 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ej265_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16. Amino acids 11 to 23 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 24, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ej265_4 should be approximately 1200 bp.

The nucleotide sequence disclosed herein for ej265_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ej265_4 demonstrated at least some similarity with sequences identified as D79053 (Human placenta cDNA 5'-end GEN-530B12), H63156 (yr50c03.r1

Homo sapiens cDNA clone 208708 5'), H64584 (yu14a12.r1 Homo sapiens cDNA clone 233758 5'), and T49682 (ya78f10.r1 Homo sapiens cDNA clone 67819 5'). The predicted amino acid sequence disclosed herein for ej265_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ej265_4 protein demonstrated at least some similarity to sequences identified as endothelial leukocyte adhesion molecule 1. Based upon sequence similarity, ej265_4 proteins and each similar protein or peptide may share at least some activity.

Clone "ey29_8"

A polynucleotide of the present invention has been identified as clone "ey29_8". ey29_8 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ey29_8 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ey29_8 protein").

The nucleotide sequence of ey29_8 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ey29_8 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Amino acids 47 to 59 are a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 60.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ey29_8 should be approximately 4000 bp.

The nucleotide sequence disclosed herein for ey29_8 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ey29_8 demonstrated at least some similarity with sequences identified as AA262521 (zs17b02.r1 Soares NbHTGBC Homo sapiens cDNA clone 685419 5'), AA429923 (zw66g01.s1 Soares testis NHT Homo sapiens cDNA clone 781200 3'), AA446080 (zw66g03.r1 Soares testis NHT Homo sapiens cDNA clone 781204 5'), F07905 (H. sapiens partial cDNA sequence; clone c-2lb06), U25125 (Gallus gallus preprogastrin gene, complete cds), W92743 (zd92g06.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 356986 3'), and Z44092 (H. sapiens partial cDNA sequence;

clone c-1sd04). Based upon sequence similarity, ey29_8 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the ey29_8 protein sequence, one centered around amino acid 120 and another around amino acid 410 of SEQ ID NO:18.

Clone "gm114_10"

A polynucleotide of the present invention has been identified as clone "gm114_10". gm114_10 was isolated from a human adult uterus cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. gm114_10 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "gm114_10 protein").

The nucleotide sequence of gm114_10 as presently determined is reported in SEQ ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the gm114_10 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone gm114_10 should be approximately 4000 bp.

The nucleotide sequence disclosed herein for gm114_10 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. gm114_10 demonstrated at least some similarity with sequences identified as AC002350 (Homo sapiens; HTGS phase 1, 46 unordered pieces), H96041 (yw61b08.r1 Soares placenta 8to9weeks 2NbHP8to9W Homo sapiens cDNA clone 256695 5'), L02529 (Rattus norvegicus Drosophila polarity gene (frizzled) homologue mRNA, complete cds), N70776 (za72g04.s1 Homo sapiens cDNA clone 298134 3'), N96041, N92163 (yz89b04.r1 Homo sapiens cDNA clone 290191 5'), U20865 (Saccharomyces cerevisiae chromosome XII cosmid 9672), and W93041 (zd93e07.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 357060 3'. The predicted amino acid sequence disclosed herein for gm114_10 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted gm114_10 protein demonstrated at least some similarity to sequences identified as U20865 (chromosome XII cosmid 9672 [Saccharomyces cerevisiae], similar to C. elegans hypothetical protein

C34E10.2 (GenBank accession number U10402)). Based upon sequence similarity, gm114_10 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the gm114_10 protein sequence centered around amino acid 150 of SEQ ID NO:20.

Deposit of Clones

Clones as294_3, aw92_1, bd316_2, bk130_4, bv131_5, bv227_1, cd265_11, ej265_4, ey29_8, and gm114_10 were deposited on June 3, 1997 with the American Type Culture Collection (10801 University Boulevard, Manassas, Virginia 20110-2209 U.S.A.) as an original deposit under the Budapest Treaty and were given the accession number ATCC 98444, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b), and the term of the deposit will comply with 37 C.F.R. § 1.806.

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Figures 1A and 1B, respectively. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of an

oligonucleotide probe that was used to isolate or to sequence each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
as294_3	SEQ ID NO:21
aw92_1	SEQ ID NO:22
bd316_2	SEQ ID NO:23
bk130_4	SEQ ID NO:24
bv131_5	SEQ ID NO:25
bv227_1	SEQ ID NO:26
cd265_11	SEQ ID NO:27
ej265_4	SEQ ID NO:28
ey29_8	SEQ ID NO:29
gm114_10	SEQ ID NO:30

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to

the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* 15(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* 62(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* 58: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided.

Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that

shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus, Ovis aries, Sus scrofa,* and *Equus caballus,* for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species (O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* 22: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* 3:103-112; Johansson *et al.*, 1995, *Genomics* 25: 682-690; Lyons *et al.*, 1997, *Nature Genetics* 15: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* 13(10): 393-399; Carver and Stubbs, 1997, *Genome Research* 7:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90%

identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides that hybridize under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, those described in Gyuris *et al.*, 1993, *Cell* 75: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* 94: 8405-8410, all of which are incorporated by reference herein) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine

levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is

evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 32

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1755 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGTGGAGTC TGTACTGGCT GCGGGGGACC CTGCTCATTT GAAAATCTGA CATCAGCTGG 60

GCAGTCGCCC CCCTCCTCCT TTCCTCCCTC TACTCTGACA CAGCACTTAG CACCTGAATC 120

TTCGTTTCTC TCCCAGGGAC CCTCCATTTT CCATATCCAG GAAAATGTGA TGCGCCACAG 180

GTATCAGCGT CTGGATCGCC ACTTCACGTT TTAGCCACAA GTGACTCAGT GGAAGATCCA 240

GAGTCAACAG AGGCTCGTCA GGAAGATGTC TACAGAAAAG GTAGACCAAA AGGAGGAAGC 300

TGGGGAAAAA GAGGTGTGCG GAGACCAGAT CAARGGACCG GACAAAGAGG AGGAACCACC 360

AGCTGCTGCA TCCCATGGCC AGGGGTGGCG TCCAGGTGGC AGAGCAGCTA GGAACGCAAG 420

GCCTGAACCT GGGGCCAGAC ACCCTGCTCT CCCGGCCATG GTCAACGACC CTCCAGTACC 480

TGCCTTACTG TGGGCCCAGG AGGTGGGCCA AGTCTTGGCA GGCCGTGCCC GCAGGCTGCT 540

GCTGCAGTTT GGGGTGCTCT TCTGCACCAT CCTCCTTTTG CTCTGGGTGT CTGTCTTCCT 600

CTATGGCTCC TTCTACTATT CCTATATGCC GACAGTCAGC CACCTCAGCC CTGTGCATTT 660

CTACTACAGG ACCGACTGTG ATTCCTCCAC CACCTCACTC TGCTCCTTCC CTGTTGCCAA 720

TGTCTCGCTG ACTAAGGGTG GACGTGATCG GGTGCTGATG TATGGACAGC CGTATCGTGT 780

TACCTTAGAG CTTGAGCTGC CAGAGTCCCC TGTGAATCAA GATTTGGGCA TGTTCTTGGT 840

CACCATTTCC TGCTACACCA GAGGTGGCCG AATCATCTCC ACTTCTTCGC GTTCGGTGAT 900

GCTGCATTAC CGCTCAGACC TGCTCCAGAT GCTGGACACA CTGGTCTTCT CTAGCCTCCT 960

GCTATTTGGC TTTGCAGAGC AGAAGCAGCT GCTGGAGGTG GAACTCTACG CAGACTATAG 1020

AGAGAACTCG TACGTGCCGA CCACTGGAGC GATCATTGAG ATCCACAGCA AGCGCATCCA 1080

GCTGTATGGA GCCTACCTCC GCATCCACGC GCACTTCACT GGGCTCAGAT ACCTGCTATA 1140

CAACTTCCCG ATGACCTGCG CCTTCATAGG TGTTGCCAGC AACTTCACCT TCCTCAGCGT 1200

CATCGTGCTC TTCAGCTACA TGCAGTGGGT GTGGGGGGGC ATCTGGCCCC GACACCGCTT 1260

CTCTTTGCAG GTTAACATCC GAAAAAGAGA CAATTCCCGG AAGGAAGTCC AACGAAGGAT 1320

CTCTGCTCAT CAGCCAGGGC CTGAAGGCCA GGAGGAGTCA ACTCCGCAAT CAGATGTTAC 1380

AGAGGATGGT GAGAGCCCTG AAGATCCCTC AGGGACAGAG GGTCAGCTGT CCGAGGAGGA 1440

GAAACCAGAT CAGCAGCCCC TGAGCGGAGA AGAGGAGCTA GAGCCTGAGG CCAGTGATGG 1500

TTCAGGCTCC TGGGAAGATG CAGCTTTGCT GACGGAGGCC AACCTGCCTG CTCCTGCTCC 1560

TGCTTCTGCT TCTGCCCCTG TCCTAGAGAC TCTGGGCAGC TCTGAACCTG CTGGGGGTGC 1620

TCTCCGACAG CGCCCCACCT GCTCTAGTTC CTGAAGAAAA GGGGCAGACT CCTCACATTC 1680

CAGCACTTTC CCACCTGACT CCTCTCCCCT CGTTTTTCCT TCAATAAACT ATTTTGTGTC 1740

AAAAAAAAAA AAAAA 1755

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 462 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Thr Glu Lys Val Asp Gln Lys Glu Glu Ala Gly Glu Lys Glu
1 5 10 15

Val Cys Gly Asp Gln Ile Lys Gly Pro Asp Lys Glu Glu Glu Pro Pro
20 25 30

Ala Ala Ala Ser His Gly Gln Gly Trp Arg Pro Gly Gly Arg Ala Ala
35 40 45

Arg Asn Ala Arg Pro Glu Pro Gly Ala Arg His Pro Ala Leu Pro Ala
50 55 60

Met Val Asn Asp Pro Pro Val Pro Ala Leu Leu Trp Ala Gln Glu Val
65 70 75 80

Gly Gln Val Leu Ala Gly Arg Ala Arg Arg Leu Leu Leu Gln Phe Gly
85 90 95

Val Leu Phe Cys Thr Ile Leu Leu Leu Leu Trp Val Ser Val Phe Leu
100 105 110

Tyr Gly Ser Phe Tyr Tyr Ser Tyr Met Pro Thr Val Ser His Leu Ser
115 120 125

Pro Val His Phe Tyr Tyr Arg Thr Asp Cys Asp Ser Ser Thr Thr Ser
130 135 140

Leu Cys Ser Phe Pro Val Ala Asn Val Ser Leu Thr Lys Gly Gly Arg

```
145                 150                 155                 160

Asp Arg Val Leu Met Tyr Gly Gln Pro Tyr Arg Val Thr Leu Glu Leu
                165                 170                 175

Glu Leu Pro Glu Ser Pro Val Asn Gln Asp Leu Gly Met Phe Leu Val
            180                 185                 190

Thr Ile Ser Cys Tyr Thr Arg Gly Gly Arg Ile Ile Ser Thr Ser Ser
        195                 200                 205

Arg Ser Val Met Leu His Tyr Arg Ser Asp Leu Leu Gln Met Leu Asp
    210                 215                 220

Thr Leu Val Phe Ser Ser Leu Leu Leu Phe Gly Phe Ala Glu Gln Lys
225                 230                 235                 240

Gln Leu Leu Glu Val Glu Leu Tyr Ala Asp Tyr Arg Glu Asn Ser Tyr
                245                 250                 255

Val Pro Thr Thr Gly Ala Ile Ile Glu Ile His Ser Lys Arg Ile Gln
            260                 265                 270

Leu Tyr Gly Ala Tyr Leu Arg Ile His Ala His Phe Thr Gly Leu Arg
        275                 280                 285

Tyr Leu Leu Tyr Asn Phe Pro Met Thr Cys Ala Phe Ile Gly Val Ala
    290                 295                 300

Ser Asn Phe Thr Phe Leu Ser Val Ile Val Leu Phe Ser Tyr Met Gln
305                 310                 315                 320

Trp Val Trp Gly Gly Ile Trp Pro Arg His Arg Phe Ser Leu Gln Val
                325                 330                 335

Asn Ile Arg Lys Arg Asp Asn Ser Arg Lys Glu Val Gln Arg Arg Ile
            340                 345                 350

Ser Ala His Gln Pro Gly Pro Glu Gly Gln Glu Glu Ser Thr Pro Gln
        355                 360                 365

Ser Asp Val Thr Glu Asp Gly Glu Ser Pro Glu Asp Pro Ser Gly Thr
    370                 375                 380

Glu Gly Gln Leu Ser Glu Glu Glu Lys Pro Asp Gln Gln Pro Leu Ser
385                 390                 395                 400

Gly Glu Glu Glu Leu Glu Pro Glu Ala Ser Asp Gly Ser Gly Ser Trp
                405                 410                 415

Glu Asp Ala Ala Leu Leu Thr Glu Ala Asn Leu Pro Ala Pro Ala Pro
            420                 425                 430

Ala Ser Ala Ser Ala Pro Val Leu Glu Thr Leu Gly Ser Ser Glu Pro
        435                 440                 445
```

Ala Gly Gly Ala Leu Arg Gln Arg Pro Thr Cys Ser Ser Ser
450 455 460

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGAATAGAGG ATTTCAAAAA GCATGCGTTT TTTGAAGGTC TAAATTGGGA AAATATACGA 60

AACCTAGAAG CACCTTATAT TCCTGATGTG AGCAGTCCCT CTGACACATC CAACTTCGAC 120

GTGGATGACG ACGTGCTGAG AAACACGGAA ATATTACCTC CTGGTTCTCA CACAGGCTTT 180

TCTGGATTAC ATTTGCCATT CATTGGTTTT ACATTCACAA CGGAAAGCTG TTTTTCTGAT 240

CGAGGCTCTC TGAAGAGCAT AATGCAGTCC AACACATTAA CCAAAGATGA GGATGTGCAG 300

CGGGACCTGG AGCACAGCCT GCAGATGGAA GCTTACGAGA GGAGGATTCG GAGGCTGGAA 360

CAGGAGAAGC TGGAGCTGAG CAGGAAGCTG CAAGAGTCCA CCCAGACCGT GCAGTCCCTC 420

CACGGCTCAT CTCGGGCCCT CAGCAATTCA AACCGAGATA AAGAAATCAA AAAGCTAAAT 480

GAAGAAATCG AACGCTTGAA GAATAAAATA GCAGATTCAA ACAGGCTGGA GCGACAGCTT 540

GAGGACACAG TGGCGCTTCG CCAAGAGCGT GAGGACTCCA CGCAGCGGCT GCGGGGGCTG 600

GAGAAGCAGC ACCGCGTGGT CCGGCAGGAG AAGGAGGAGC TGCACAAGCA ACTGGTTGAA 660

GCCTCAGAGC GGTTGAAATC CCAGGCCAAG GAACTCAAAG ATGCCCATCA GCAGCGAAAG 720

CTGGCCCTGC AGGAGTTCTC GGAGCTGAAC GAGCGCATGG CAGAGCTCCG TGCCCAGAAG 780

CAGAAGGTGT CCCGGCAGCT GCGAGACAAG GAGGAGGAGA TGGAGGTGGC CACGCAGAAG 840

GTGGACGCCA TGCGGCAGGA AATGCGGAGA GCTGAGAAGC TCAGGAAAGA GCTGGAAGCT 900

CAGCTTGATG ATGCTGTTGC TGAGGCCTCC AAGGAGCGCA AGCTTCGTGA GCACAGCGAG 960

AACTTCTGCA AGCAAATGGA AAGCGAGCTG GAGGCCCTCA AGGTGAAGCA AGGAGGCCGG 1020

GGAGCGGGTG CCACCTTAGA GCACCAGCAA GAGATTTCCA AAATCAAATC CGAGCTGGAG 1080

AAGAAAGTCT TATTTTATGA AGAGGAATTG GTCAGACGTG AGGCCTCCCA TGTGCTAGAA 1140

GTGAAAAATG TGAAGAAGGA GGTGCATGAT TCAGAAAGCC ACCAGCTGGC CCTGCAGAAA 1200

GAAATCTTGA TGTTAAAAGA TAAGTTAGAA AAGTCAAAGC GAGAACGGCA TAACGAGATG 1260

GAGGAGGCAG TAGGTACAAT AAAAGATAAA TACGAACGAG AAAGAGCGAT GCTGTTTGAT 1320

GAAAACAAGA AGCTAACTGC TGAAAATGAA AAGCTCTGTT CCTTTGTGGA TAAACTCACA 1380

GCTCAAAATA GACAGCTGGA GGATGAGCTG CAGGATCTGG CAGCCAAGAA GGAGTCAGTG 1440

GCCCACTGGG AAGCTCAGAT TGCGGAAATC ATTCAGTGGG TCAGTGACGA GAAAGATGCC 1500

CGGGGTTACC TTCAAGCTCT TGCTTCCAAG ATGACCGAAG AGCTCGAGGC TTTGAGGAGT 1560

TCTAGTCTGG GGTCAAGAAC ACTGGACCCG CTGTGGAAGG TGCGCCGCAG CCAGAAGCTG 1620

GACATGTCCG CGCGGCTGGA GCTGCAGTCG GCCCTGGAGG CGGAGATCCG GGCCAAGCAG 1680

CTTGTCCAGG AGGAGCTCAG GAAGGTCAAG GACGCCAACC TCACCTTGGA AAGCAAACYA 1740

AWGGATTCCG AAGCCAAAAA CAGAGAATTA TTAGAAGAAA TGGAAATTTT GAAGAAAAAG 1800

ATGGAAGAAA AATTCAGAGC AGATACTGGG CTCAAACTTC CAGATTTTCA GGATTCCATT 1860

TTTGAGTATT TCAACACTGC TCCTCTTGCA CATGACCTGA CATTTAGAAC CAGCTCAGCT 1920

AGTGAGCAAG AAACACAAGC TCCGAAGCCA GAAGCGTCCC CGTCGATGTC TGTGGCTGCA 1980

TCAGAGCAGC AGGAGGACAT GGCTCGGCCC CCGCAGAGGC CATCCGCTGT GCCGTTGCCC 2040

ACCACGCAGG CCCTGGCTCT GGCTGGACCG AAGCCAAAAG CTCACCAGTT CAGCATCAAG 2100

TCCTTCTCCA GCCCTACTCA GTGCAGCCAC TGCACCTCCC TGATGGTTGG GCTGATCCGG 2160

CAGGGCTACG CCTGCGAGGT GTGTTCCTTT GCTTGCCACG TGTCCTGCAA AGACGGTGCC 2220

CCCCAGGTGT GCCCAATACC TCCCGAGCAG TCCAAGAGGC CTCTGGGCGT GGACGTGCAG 2280

CGAGGCATCG GAACAGCCTA CAAAGGCCAT GTCAAGGTCC CAAAGCCCAC GGGGGTGAAG 2340

AAGGGATGGC AGCGCGCATA TGCAGTCGTC TGTGACTGCA AGCTCTTCCT GTATGATCTG 2400

CCTGAAGGAA AATCCACCCA GCCTGGTGTC ATTGCGAGCC AAGTCTTGGA TCTCAGAGAT 2460

GACGAGTTTT CCGTGAGCTC AGTCCTGGCC TCAGATGTCA TTCATGCTAC ACGCCGAGAT 2520

ATTCCATGTA TATTCAGGGT GACGGCCTCT CTCTTAGGTG CACCTTCTAA GACCAGCTCG 2580

CTGCTCATTC TGACAGAAAA TGAGAATGAA AAGAGGAAGT GGGTTGGGAT TCTAGAAGGA 2640

CTCCAGTCCA TCCTTCATAA AAACCGGCTG AGGAATCAGG TCGTGCATGT TCCCTTGGAA 2700

GCCTACGACA GCTCGCTGCC TCTCATCAAG GCCATCCTGA CAGCTGCCAT CGTGGATGCA 2760

GACAGGATTG CAGTCGGCCT AGAAGAAGGG CTCTATGTCA TAGAGGTCAC CCGAGATGTG 2820

ATCGTCCGTG CCGCTGACTG TAAGAAGGTA CACCAGATCG AGCTTGCTCC CAGGGAGAAG 2880

ATCGTAATCC TCCTCTGTGG CCGGAACCAC CATGTGCACC TCTATCCGTG GTCGTCCCTT 2940

GATGGAGCGG AAGGCAGCTT TGACATCAAG CTTCCGGAAA CCAAAGGCTG CCAGCTCATG 3000

GCCACGGCCA CACTCAAGAG GARCTCTGGC ACCTGCCTGT TTGTGGCCGT GAAACGGCTG 3060

ATCCTTTGCT ATGAGATCCA GAAAATAAAG CCATATTGAA TGATAAAAAA AAAAAAAAAA 3120

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 3180

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA 3213

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 945 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gln Ser Asn Thr Leu Thr Lys Asp Glu Asp Val Gln Arg Asp Leu
1 5 10 15

Glu His Ser Leu Gln Met Glu Ala Tyr Glu Arg Arg Ile Arg Arg Leu
20 25 30

Glu Gln Glu Lys Leu Glu Leu Ser Arg Lys Leu Gln Glu Ser Thr Gln
35 40 45

Thr Val Gln Ser Leu His Gly Ser Ser Arg Ala Leu Ser Asn Ser Asn
50 55 60

Arg Asp Lys Glu Ile Lys Lys Leu Asn Glu Glu Ile Glu Arg Leu Lys
65 70 75 80

Asn Lys Ile Ala Asp Ser Asn Arg Leu Glu Arg Gln Leu Glu Asp Thr
85 90 95

Val Ala Leu Arg Gln Glu Arg Glu Asp Ser Thr Gln Arg Leu Arg Gly
100 105 110

Leu Glu Lys Gln His Arg Val Val Arg Gln Glu Lys Glu Glu Leu His
115 120 125

Lys Gln Leu Val Glu Ala Ser Glu Arg Leu Lys Ser Gln Ala Lys Glu
130 135 140

```
Leu Lys Asp Ala His Gln Gln Arg Lys Leu Ala Leu Gln Glu Phe Ser
145                 150                 155                 160

Glu Leu Asn Glu Arg Met Ala Glu Leu Arg Ala Gln Lys Gln Lys Val
                165                 170                 175

Ser Arg Gln Leu Arg Asp Lys Glu Glu Glu Met Glu Val Ala Thr Gln
            180                 185                 190

Lys Val Asp Ala Met Arg Gln Glu Met Arg Arg Ala Glu Lys Leu Arg
        195                 200                 205

Lys Glu Leu Glu Ala Gln Leu Asp Asp Ala Val Ala Glu Ala Ser Lys
    210                 215                 220

Glu Arg Lys Leu Arg Glu His Ser Glu Asn Phe Cys Lys Gln Met Glu
225                 230                 235                 240

Ser Glu Leu Glu Ala Leu Lys Val Lys Gln Gly Gly Arg Gly Ala Gly
                245                 250                 255

Ala Thr Leu Glu His Gln Gln Glu Ile Ser Lys Ile Lys Ser Glu Leu
            260                 265                 270

Glu Lys Lys Val Leu Phe Tyr Glu Glu Glu Leu Val Arg Arg Glu Ala
        275                 280                 285

Ser His Val Leu Glu Val Lys Asn Val Lys Lys Glu Val His Asp Ser
    290                 295                 300

Glu Ser His Gln Leu Ala Leu Gln Lys Glu Ile Leu Met Leu Lys Asp
305                 310                 315                 320

Lys Leu Glu Lys Ser Lys Arg Glu Arg His Asn Glu Met Glu Glu Ala
                325                 330                 335

Val Gly Thr Ile Lys Asp Lys Tyr Glu Arg Glu Arg Ala Met Leu Phe
            340                 345                 350

Asp Glu Asn Lys Lys Leu Thr Ala Glu Asn Glu Lys Leu Cys Ser Phe
        355                 360                 365

Val Asp Lys Leu Thr Ala Gln Asn Arg Gln Leu Glu Asp Glu Leu Gln
    370                 375                 380

Asp Leu Ala Ala Lys Lys Glu Ser Val Ala His Trp Glu Ala Gln Ile
385                 390                 395                 400

Ala Glu Ile Ile Gln Trp Val Ser Asp Glu Lys Asp Ala Arg Gly Tyr
                405                 410                 415

Leu Gln Ala Leu Ala Ser Lys Met Thr Glu Glu Leu Glu Ala Leu Arg
            420                 425                 430

Ser Ser Ser Leu Gly Ser Arg Thr Leu Asp Pro Leu Trp Lys Val Arg
```

435 440 445

Arg Ser Gln Lys Leu Asp Met Ser Ala Arg Leu Glu Leu Gln Ser Ala
450 455 460

Leu Glu Ala Glu Ile Arg Ala Lys Gln Leu Val Gln Glu Glu Leu Arg
465 470 475 480

Lys Val Lys Asp Ala Asn Leu Thr Leu Glu Ser Lys Xaa Xaa Asp Ser
485 490 495

Glu Ala Lys Asn Arg Glu Leu Leu Glu Glu Met Glu Ile Leu Lys Lys
500 505 510

Lys Met Glu Glu Lys Phe Arg Ala Asp Thr Gly Leu Lys Leu Pro Asp
515 520 525

Phe Gln Asp Ser Ile Phe Glu Tyr Phe Asn Thr Ala Pro Leu Ala His
530 535 540

Asp Leu Thr Phe Arg Thr Ser Ser Ala Ser Glu Gln Glu Thr Gln Ala
545 550 555 560

Pro Lys Pro Glu Ala Ser Pro Ser Met Ser Val Ala Ala Ser Glu Gln
565 570 575

Gln Glu Asp Met Ala Arg Pro Pro Gln Arg Pro Ser Ala Val Pro Leu
580 585 590

Pro Thr Thr Gln Ala Leu Ala Leu Ala Gly Pro Lys Pro Lys Ala His
595 600 605

Gln Phe Ser Ile Lys Ser Phe Ser Ser Pro Thr Gln Cys Ser His Cys
610 615 620

Thr Ser Leu Met Val Gly Leu Ile Arg Gln Gly Tyr Ala Cys Glu Val
625 630 635 640

Cys Ser Phe Ala Cys His Val Ser Cys Lys Asp Gly Ala Pro Gln Val
645 650 655

Cys Pro Ile Pro Pro Glu Gln Ser Lys Arg Pro Leu Gly Val Asp Val
660 665 670

Gln Arg Gly Ile Gly Thr Ala Tyr Lys Gly His Val Lys Val Pro Lys
675 680 685

Pro Thr Gly Val Lys Lys Gly Trp Gln Arg Ala Tyr Ala Val Val Cys
690 695 700

Asp Cys Lys Leu Phe Leu Tyr Asp Leu Pro Glu Gly Lys Ser Thr Gln
705 710 715 720

Pro Gly Val Ile Ala Ser Gln Val Leu Asp Leu Arg Asp Asp Glu Phe
725 730 735

```
Ser Val Ser Ser Val Leu Ala Ser Asp Val Ile His Ala Thr Arg Arg
            740                 745                 750

Asp Ile Pro Cys Ile Phe Arg Val Thr Ala Ser Leu Leu Gly Ala Pro
        755                 760                 765

Ser Lys Thr Ser Ser Leu Leu Ile Leu Thr Glu Asn Glu Asn Glu Lys
    770                 775                 780

Arg Lys Trp Val Gly Ile Leu Glu Gly Leu Gln Ser Ile Leu His Lys
785                 790                 795                 800

Asn Arg Leu Arg Asn Gln Val Val His Val Pro Leu Glu Ala Tyr Asp
                805                 810                 815

Ser Ser Leu Pro Leu Ile Lys Ala Ile Leu Thr Ala Ala Ile Val Asp
            820                 825                 830

Ala Asp Arg Ile Ala Val Gly Leu Glu Glu Gly Leu Tyr Val Ile Glu
        835                 840                 845

Val Thr Arg Asp Val Ile Val Arg Ala Ala Asp Cys Lys Lys Val His
    850                 855                 860

Gln Ile Glu Leu Ala Pro Arg Glu Lys Ile Val Ile Leu Leu Cys Gly
865                 870                 875                 880

Arg Asn His His Val His Leu Tyr Pro Trp Ser Ser Leu Asp Gly Ala
                885                 890                 895

Glu Gly Ser Phe Asp Ile Lys Leu Pro Glu Thr Lys Gly Cys Gln Leu
            900                 905                 910

Met Ala Thr Ala Thr Leu Lys Arg Xaa Ser Gly Thr Cys Leu Phe Val
        915                 920                 925

Ala Val Lys Arg Leu Ile Leu Cys Tyr Glu Ile Gln Lys Ile Lys Pro
    930                 935                 940

Tyr
945
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1315 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAGGGCACTT AATCCCAATG AACTGTATGC TTAAAAATAA TTTAAATGAT AAACTTTGTG 60

TTATGTATAC TTTACCACAA TAAGAAAAAG TATTTTAGTA CTAGTGGTAA ATAGTTTTTA 120

TTTAATAGAC TTATATTTTA AAGCTTAAAA ATAATTTAGC TTCTAGAGTA TTACGTTTTT 180

CTTCATGGGA ACTTCAAAAA GCAAGTCACT AAATCCAAGA ATTTTAAAGA AAAAACCCAA 240

ATACATGATT TATGCTGCAT CTGGTATAGA TTTTTAAAAG ACTAGTCAAT CTAAGCTCTA 300

AACTATTAAA TGACAAACCA TTTCATATGT CATTGCATAT TCCTATGTAC CACATTCTCA 360

TATTTCTGTT ATGGGCATGA AGGGGTGTTT GATGCTTCCA TGCCATAATA ACCATGACTA 420

TCACAACCAT TGAAATAAAG GTTCTTGCAG TATTTTCAGG ATGGTCCCAG AAATTTAAAT 480

TAATCTCTCA TCCATTGGCT TTTGCTACTT TAGGTTAATA TTAAAATATA ACATACATTT 540

TTGGGGTTTA TGCTGTTAGC TCCAAACCAA AAGATTTTGG AAATTTATTT TGGAAATTTT 600

GTGTTTAGAA TATGAATAAA TCTGCTTATT CAGAAAAATT AAACCTTGAT AACTTGGGAC 660

CTCCTATTCC TGTATGTTCT CTGACATACA TTGAGGGATT TGGCTCTCTT TTGTTTATTT 720

GTTTTACTAG TCAGACATTC CTTTGGCTGC CCATACTTAA TTCTGTTGGG TGTTTCCGCC 780

CCCGCCCTCA GCTTCTGCAG CTACTCTGAT CAACATCCGC AATGCCAGGA AACACTTTGA 840

AAAGCTGGAA AGAGTGGATG GACCAAAGCA GTGTCTTCTC ATGCGCTAAA CATTGATGAA 900

TATTGTTTCA CACAAAAATT AAAAGTTTCC TAATTAATGT TGTATTCATA TATGTAGGCT 960

CTGAAATGTT GTGATGCTTA TTGCTTCTGT ATTTCTTCTC TACTCCCTAG TCTTAATGTT 1020

TAACCTTGAA TGCTATTAAC TTAAATAGCC ATTGAGGAGT TAGAAGATGA ATTGTTCATG 1080

AAGTCGGTGT TACATAAAAG TAGGTGATAT GTAAGTTTTC TGATAACAAG GTTCTAATAG 1140

TGTTTAAATG TACTGGTAAC CTGGTTCCAA TAGTTGTGTT TGCCCAAGCC TTTCTCGGCA 1200

TCATCTTGTA TTCCTTATCA GATAGTAAGT AACCTGTAAG TTTGGAGTAT TACTGTTTTC 1260

TCAGCATGCA TTAAAAATAT TCCTTAACTT CAATTGTAAA AAAAAAAAAA AAAAA 1315

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Lys Ser Ala Tyr Ser Glu Lys Leu Asn Leu Asp Asn Leu Gly
1 5 10 15

Pro Pro Ile Pro Val Cys Ser Leu Thr Tyr Ile Glu Gly Phe Gly Ser
20 25 30

Leu Leu Phe Ile Cys Phe Thr Ser Gln Thr Phe Leu Trp Leu Pro Ile
35 40 45

Leu Asn Ser Val Gly Cys Phe Arg Pro Arg Pro Gln Leu Leu Gln Leu
50 55 60

Leu
65

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 519 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TAGGCCATGA AGGCCGAATC GGCCTTCATG GCCTACGCTT ACACAATACC CACCATGTCC 60

CAGGCTGGTG CTCAGGAAGC CCCTATCAAG AAGAAGCGCC CCCCTGTGAA GGAGGAGGAC 120

CTGAAGGGGG CCCGAGGAAA CCTGACCAAG AACCAGGAAA TCAAGTCCAA GACCTACCAG 180

GTCATGCGAG AGTGTGAGCA AGCTGGCTCG GCCGCCCCGT CGGTGTTCAG CCGCACCCGC 240

ACAGGTACCG AGACTGTCTT TGAGAAGCCC AAAGCCGGAC CCACCAAGAG TGTCTTCGGC 300

TGAGAAGTGT GCGCCACTCC CCTTGCTGCC CGAATGCTCG GAAACAGGAG CCTTACCCAG 360

GAACTCTTTT TTATGCCAGA ACGCTTCCTC TCCCCTGCTG TCTCTGGGGC TGCCACCCTC 420

CCCCACAGTC CAGGCCCTTC AGCCAAGGGC TCTGCACCAG CACCTTGGAA GCACCAATAA 480

AGAGGATGCC CACGTGGCCC CAGCAAAAAA AAAAAAAAA 519

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Lys Ala Glu Ser Ala Phe Met Ala Tyr Ala Tyr Thr Ile Pro Thr
1 5 10 15

Met Ser Gln Ala Gly Ala Gln Glu Ala Pro Ile Lys Lys Lys Arg Pro
20 25 30

Pro Val Lys Glu Glu Asp Leu Lys Gly Ala Arg Gly Asn Leu Thr Lys
35 40 45

Asn Gln Glu Ile Lys Ser Lys Thr Tyr Gln Val Met Arg Glu Cys Glu
50 55 60

Gln Ala Gly Ser Ala Ala Pro Ser Val Phe Ser Arg Thr Arg Thr Gly
65 70 75 80

Thr Glu Thr Val Phe Glu Lys Pro Lys Ala Gly Pro Thr Lys Ser Val
85 90 95

Phe Gly

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2788 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACGGCGACC AAACCCAGCT AGGTCAGACG AGAAAGATAA AAACTCTCCA GATGTCTTCC 60

AGTAATGTCG AAGTTTTTAT CCCAGTGTCA CAAGGAAACA CCAATGGCTT CCCCGCGACA 120

GCTTCCAATG ACCTGAAGGC ATTTACTGAA GGAGCTGTGT TAAGTTTTCA TAACATCTGC 180

TATCGAGTAA AACTGAAGAG TGGCTTTCTA CCTTGTCGAA AACCAGTTGA GAAAGAAATA 240

TTATCGAATA TCAATGGGAT CATGAAACCT GGTCTCAACG CCATCCTGGG ACCCACAGGT 300

GGARGCAAAT CTTCGTTATT AGATGTCTTA GCTGCAAGGA AAGATCCAAG TGGATTATCT 360

GGAGATGTTC TGATAAATGG AGCACCGCGA CCTGCCAATT TCAAATGTAA TTCAGGTTAC 420

GTGGTACAAG TTGGAACTCA GTTTATCCGT GGTGTGTCTG GAGGAGAAAG AAAAAGGACT 480

AGTATAGGAA TGGAGCTTAT CACTGATCCT TCCATCTTGT TCTTGGATGA GCCTACAACT 540

GGCTTAGACT CAAGCACAGC AAATGCTGTC CTTTTGCTCC TGAAAAGGAT GTCTAAGCAG 600

GGACGAACAA TCATCTTCTC CATTCATCAG CCTCGATATT CCATCTTCAA GTTGTTTGAT 660

AGCCTCACCT TATTGGCCTC AGGAAGACTT ATGTTCCACG GGCCTGCTCA GGAGGCCTTG 720

GGATACTTTG AATCAGCTGG TTATCACTGT GAGGCCTATA ATAACCCTGC AGACTTCTTC 780

TTGGACATCA TTAATGGAGA TTCCACTGCT GTGGCATTAA ACAGAGAAGA AGACTTTAAA 840

GCCACAGAGA TCATAGAGCC TTCCAAGCAG GATAAGCCAC TCATAGAAAA ATTAGCGGAG 900

ATTTATGTCA ACTCCTGCTT CTACAAAGAG ACAAAAGCTG AATTACATCA ACTTTCCGGG 960

GGTGAGAAGA AGAAGAAGAT CACAGTCTTC AAGGAGATCA GCTACACCAC CTCCTTCTGT 1020

CATCAACTCA GATGGGTTTC CAAGCGTTCA TTCAAAAACT TGCTGGGTAA TCCCCAGGCC 1080

TCTATAGCTC AGATCATTGT CACAGTCGTA CTGGGACTGG TTATAGGTGC CATTTACTTT 1140

GGGCTAAAAA ATGATTCTAC TGGAATCCAG AACAGAGCTG GGGTTCTCTT CTTCCTGACG 1200

ACCAACCAGT GTTTCAGCAG TGTTTCAGCC GTGGAACTCT TTGTGGTAGA GAAGAAGCTC 1260

TTCATACATG AATACATCAG CGGATACTAC AGAGTGTCAT CTTATTTCCT TGGAAAACTG 1320

TTATCTGATT TATTACCCAT GAGGATGTTA CCAAGTATTA TATTTACCTG TATAGTGTAC 1380

TTCATGTTAG GATTGAAGCC AAAGGCAGAT GCCTTCTTCG TTATGATGTT TACCCTTATG 1440

ATGGTGGCTT ATTCAGCCAG TTCCATGGCA CTGGCCATAG CAGCAGGTCA GAGTGTGGTT 1500

TCTGTAGCAA CACTTCTCAT GACCATCTGT TTTGTGTTTA TGATGATTTT TTCAGGTCTG 1560

TTGGTCAATC TCACAACCAT TGCATCTTGG CTGTCATGGC TTCAGTACTT CAGCATTCCA 1620

CGATATGGAT TTACGGCTTT GCAGCATAAT GAATTTTTGG GACAAAACTT CTGCCCAGGA 1680

CTCAATGCAA CAGGAAACAA TCCTTGTAAC TATGCAACAT GTACTGGCGA AGAATATTTG 1740

GTAAAGCAGG GCATCGATCT CTCACCCTGG GGCTTGTGGA AGAATCACGT GGCCTTGGCT 1800

TGTATGATTG TTATTTTCCT CACAATTGCC TACCTGAAAT TGTTATTTCT TAAAAAATAT 1860

TCTTAAATTT CCCCTTAATT CAGTATGATT TATCCTCACA TAAAAAAGAA GCACTTTGAT 1920

TGAAGTATTC AATCAAGTTT TTTTGGTTGT TTTCTGTTCC CTTGCCATCA CACTGTTGCA 1980

CAGCAGCAAT TGTTTTAAAG AGATACATTT TTAGAAATCA CAACAAACTG AATTAAACAT 2040

GAAAGAACCC AAGACATCAT GTATCGCATA TTAGTTAATC TCCTCAGACA GTAACCATGG 2100

GGAAGAAATC TGGTCTAATT TATTAATCTA AAAAAGGAGA ATTGAATTCT GGAAACTCCT 2160

GACAAGTTAT TACTGTCTCT GGCATTTGTT TCCTCATCTT TAAAATGAAT AGGTAGGTTA 2220

GTAGCCCTTC AGTCTTAATA CTTTATGATG CTATGGTTTG CCATTATTTA ATAAATGACA 2280

AATGTATTAA TGCTAAAAAA AAAAAAAAAA AGCGGCCTTC ATGGCCTAGA GATTTCAACT 2340

TAACTTGACC GCTCTGAGCT AAACCTAGCC CCAAACCCAC TCCACCTTAT TACCAGACAA 2400

CCTTAACCAA ACCATTTACC CAAATAAAGT ATAGGCGATA GAAATTGAAA CCTGGCGCAA 2460

TAGATATAGT ACCGCAAGGG AAAGATGAAA AATTATAACC AAGCATAATA TAGCAAGGAC 2520

TAACCCCTAT ACCTTCTGCA TAATGAATTA ACTAGAAATA ACTTTGCAAG GAGAGCCAAA 2580

GCTAAGACCC CCGAAACCAG ACGAGCTACC TAAGAACAGC TAAAAGAGCA CACCCGTCTA 2640

TGTAGCAAAA TAGTGGGAAG ATTTATAGGT AGAGGCGACA AACCTACCGA GCCTGGTGAT 2700

AGCTGGTTGT CCCAGAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2760

AAAAAAAAAA AAAAAAAAAA AAAAAAAA 2788

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 604 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Ser Ser Asn Val Glu Val Phe Ile Pro Val Ser Gln Gly Asn
1 5 10 15

Thr Asn Gly Phe Pro Ala Thr Ala Ser Asn Asp Leu Lys Ala Phe Thr
20 25 30

Glu Gly Ala Val Leu Ser Phe His Asn Ile Cys Tyr Arg Val Lys Leu
35 40 45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GAGGGCACTT AATCCCAATG AACTGTATGC TTAAAAATAA TTTAAATGAT AAACTTTGTG 60

TTATGTATAC TTTACCACAA TAAGAAAAAG TATTTTAGTA CTAGTGGTAA ATAGTTTTTA 120

TTTAATAGAC TTATATTTTA AAGCTTAAAA ATAATTTAGC TTCTAGAGTA TTACGTTTTT 180

CTTCATGGGA ACTTCAAAAA GCAAGTCACT AAATCCAAGA ATTTTAAAGA AAAAACCCAA 240

ATACATGATT TATGCTGCAT CTGGTATAGA TTTTTAAAAG ACTAGTCAAT CTAAGCTCTA 300

AACTATTAAA TGACAAACCA TTTCATATGT CATTGCATAT TCCTATGTAC CACATTCTCA 360

TATTTCTGTT ATGGGCATGA AGGGGTGTTT GATGCTTCCA TGCCATAATA ACCATGACTA 420

TCACAACCAT TGAAATAAAG GTTCTTGCAG TATTTTCAGG ATGGTCCCAG AAATTTAAAT 480

TAATCTCTCA TCCATTGGCT TTTGCTACTT TAGGTTAATA TTAAAATATA ACATACATTT 540

TTGGGGTTTA TGCTGTTAGC TCCAAACCAA AAGATTTTGG AAATTTATTT TGGAAATTTT 600

GTGTTTAGAA TATGAATAAA TCTGCTTATT CAGAAAAATT AAACCTTGAT AACTTGGGAC 660

CTCCTATTCC TGTATGTTCT CTGACATACA TTGAGGGATT TGGCTCTCTT TTGTTTATTT 720

GTTTTACTAG TCAGACATTC CTTTGGCTGC CCATACTTAA TTCTGTTGGG TGTTTCCGCC 780

CCCGCCCTCA GCTTCTGCAG CTACTCTGAT CAACATCCGC AATGCCAGGA AACACTTTGA 840

AAAGCTGGAA AGAGTGGATG GACCAAAGCA GTGTCTTCTC ATGCGCTAAA CATTGATGAA 900

TATTGTTTCA CACAAAAATT AAAAGTTTCC TAATTAATGT TGTATTCATA TATGTAGGCT 960

CTGAAATGTT GTGATGCTTA TTGCTTCTGT ATTTCTTCTC TACTCCCTAG TCTTAATGTT 1020

TAACCTTGAA TGCTATTAAC TTAAATAGCC ATTGAGGAGT TAGAAGATGA ATTGTTCATG 1080

AAGTCGGTGT TACATAAAAG TAGGTGATAT GTAAGTTTTC TGATAACAAG GTTCTAATAG 1140

TGTTTAAATG TACTGGTAAC CTGGTTCCAA TAGTTGTGTT TGCCCAAGCC TTTCTCGGCA 1200

TCATCTTGTA TTCCTTATCA GATAGTAAGT AACCTGTAAG TTTGGAGTAT TACTGTTTTC 1260

TCAGCATGCA TTAAAAATAT TCCTTAACTT CAATTGTAAA AAAAAAAAAA AAAAA 1315

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Lys Ser Ala Tyr Ser Glu Lys Leu Asn Leu Asp Asn Leu Gly
1 5 10 15

Pro Pro Ile Pro Val Cys Ser Leu Thr Tyr Ile Glu Gly Phe Gly Ser
20 25 30

Leu Leu Phe Ile Cys Phe Thr Ser Gln Thr Phe Leu Trp Leu Pro Ile
35 40 45

Leu Asn Ser Val Gly Cys Phe Arg Pro Arg Pro Gln Leu Leu Gln Leu
50 55 60

Leu
65

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 519 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TAGGCCATGA AGGCCGAATC GGCCTTCATG GCCTACGCTT ACACAATACC CACCATGTCC 60

CAGGCTGGTG CTCAGGAAGC CCCTATCAAG AAGAAGCGCC CCCCTGTGAA GGAGGAGGAC 120

CTGAAGGGGG CCCGAGGAAA CCTGACCAAG AACCAGGAAA TCAAGTCCAA GACCTACCAG 180

GTCATGCGAG AGTGTGAGCA AGCTGGCTCG GCCGCCCCGT CGGTGTTCAG CCGCACCCGC 240

ACAGGTACCG AGACTGTCTT TGAGAAGCCC AAAGCCGGAC CCACCAAGAG TGTCTTCGGC 300

TGAGAAGTGT GCGCCACTCC CCTTGCTGCC CGAATGCTCG GAAACAGGAG CCTTACCCAG 360

GAACTCTTTT TTATGCCAGA ACGCTTCCTC TCCCCTGCTG TCTCTGGGGC TGCCACCCTC 420

CCCCACAGTC CAGGCCCTTC AGCCAAGGGC TCTGCACCAG CACCTTGGAA GCACCAATAA 480

AGAGGATGCC CACGTGGCCC CAGCAAAAAA AAAAAAAAA 519

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Lys Ala Glu Ser Ala Phe Met Ala Tyr Ala Tyr Thr Ile Pro Thr
1 5 10 15

Met Ser Gln Ala Gly Ala Gln Glu Ala Pro Ile Lys Lys Lys Arg Pro
20 25 30

Pro Val Lys Glu Glu Asp Leu Lys Gly Ala Arg Gly Asn Leu Thr Lys
35 40 45

Asn Gln Glu Ile Lys Ser Lys Thr Tyr Gln Val Met Arg Glu Cys Glu
50 55 60

Gln Ala Gly Ser Ala Ala Pro Ser Val Phe Ser Arg Thr Arg Thr Gly
65 70 75 80

Thr Glu Thr Val Phe Glu Lys Pro Lys Ala Gly Pro Thr Lys Ser Val
85 90 95

Phe Gly

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2788 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACGGCGACC AAACCCAGCT AGGTCAGACG AGAAAGATAA AAACTCTCCA GATGTCTTCC 60

AGTAATGTCG AAGTTTTTAT CCCAGTGTCA CAAGGAAACA CCAATGGCTT CCCCGCGACA 120

GCTTCCAATG ACCTGAAGGC ATTTACTGAA GGAGCTGTGT TAAGTTTTCA TAACATCTGC 180

TATCGAGTAA AACTGAAGAG TGGCTTTCTA CCTTGTCGAA AACCAGTTGA GAAAGAAATA 240

TTATCGAATA TCAATGGGAT CATGAAACCT GGTCTCAACG CCATCCTGGG ACCCACAGGT 300

GGARGCAAAT CTTCGTTATT AGATGTCTTA GCTGCAAGGA AAGATCCAAG TGGATTATCT 360

GGAGATGTTC TGATAAATGG AGCACCGCGA CCTGCCAATT TCAAATGTAA TTCAGGTTAC 420

GTGGTACAAG TTGGAACTCA GTTTATCCGT GGTGTGTCTG GAGGAGAAAG AAAAAGGACT 480

AGTATAGGAA TGGAGCTTAT CACTGATCCT TCCATCTTGT TCTTGGATGA GCCTACAACT 540

GGCTTAGACT CAAGCACAGC AAATGCTGTC CTTTTGCTCC TGAAAAGGAT GTCTAAGCAG 600

GGACGAACAA TCATCTTCTC CATTCATCAG CCTCGATATT CCATCTTCAA GTTGTTTGAT 660

AGCCTCACCT TATTGGCCTC AGGAAGACTT ATGTTCCACG GGCCTGCTCA GGAGGCCTTG 720

GGATACTTTG AATCAGCTGG TTATCACTGT GAGGCCTATA ATAACCCTGC AGACTTCTTC 780

TTGGACATCA TTAATGGAGA TTCCACTGCT GTGGCATTAA ACAGAGAAGA AGACTTTAAA 840

GCCACAGAGA TCATAGAGCC TTCCAAGCAG GATAAGCCAC TCATAGAAAA ATTAGCGGAG 900

ATTTATGTCA ACTCCTCCTT CTACAAAGAG ACAAAAGCTG AATTACATCA ACTTTCCGGG 960

GGTGAGAAGA AGAAGAAGAT CACAGTCTTC AAGGAGATCA GCTACACCAC CTCCTTCTGT 1020

CATCAACTCA GATGGGTTTC CAAGCGTTCA TTCAAAAACT TGCTGGGTAA TCCCCAGGCC 1080

TCTATAGCTC AGATCATTGT CACAGTCGTA CTGGGACTGG TTATAGGTGC CATTTACTTT 1140

GGGCTAAAAA ATGATTCTAC TGGAATCCAG AACAGAGCTG GGGTTCTCTT CTTCCTGACG 1200

ACCAACCAGT GTTTCAGCAG TGTTTCAGCC GTGGAACTCT TTGTGGTAGA GAAGAAGCTC 1260

TTCATACATG AATACATCAG CGGATACTAC AGAGTGTCAT CTTATTTCCT TGGAAAACTG 1320

TTATCTGATT TATTACCCAT GAGGATGTTA CCAAGTATTA TATTTACCTG TATAGTGTAC 1380

TTCATGTTAG GATTGAAGCC AAAGGCAGAT GCCTTCTTCG TTATGATGTT TACCCTTATG 1440

ATGGTGGCTT ATTCAGCCAG TTCCATGGCA CTGGCCATAG CAGCAGGTCA GAGTGTGGTT 1500

TCTGTAGCAA CACTTCTCAT GACCATCTGT TTTGTGTTTA TGATGATTTT TTCAGGTCTG 1560

TTGGTCAATC TCACAACCAT TGCATCTTGG CTGTCATGGC TTCAGTACTT CAGCATTCCA 1620

CGATATGGAT TTACGGCTTT GCAGCATAAT GAATTTTTGG GACAAAACTT CTGCCCAGGA 1680

CTCAATGCAA CAGGAAACAA TCCTTGTAAC TATGCAACAT GTACTGGCGA AGAATATTTG 1740

GTAAAGCAGG GCATCGATCT CTCACCCTGG GGCTTGTGGA AGAATCACGT GGCCTTGGCT 1800

TGTATGATTG TTATTTTCCT CACAATTGCC TACCTGAAAT TGTTATTTCT TAAAAAATAT 1860

TCTTAAATTT CCCCTTAATT CAGTATGATT TATCCTCACA TAAAAAAGAA GCACTTTGAT 1920

TGAAGTATTC AATCAAGTTT TTTTGGTTGT TTTCTGTTCC CTTGCCATCA CACTGTTGCA 1980

CAGCAGCAAT TGTTTTAAAG AGATACATTT TTAGAAATCA CAACAAACTG AATTAAACAT 2040

GAAAGAACCC AAGACATCAT GTATCGCATA TTAGTTAATC TCCTCAGACA GTAACCATGG 2100

GGAAGAAATC TGGTCTAATT TATTAATCTA AAAAAGGAGA ATTGAATTCT GGAAACTCCT 2160

GACAAGTTAT TACTGTCTCT GGCATTTGTT TCCTCATCTT TAAAATGAAT AGGTAGGTTA 2220

GTAGCCCTTC AGTCTTAATA CTTTATGATG CTATGGTTTG CCATTATTTA ATAAATGACA 2280

AATGTATTAA TGCTAAAAAA AAAAAAAAAA AGCGGCCTTC ATGGCCTAGA GATTTCAACT 2340

TAACTTGACC GCTCTGAGCT AAACCTAGCC CCAAACCCAC TCCACCTTAT TACCAGACAA 2400

CCTTAACCAA ACCATTTACC CAAATAAAGT ATAGGCGATA GAAATTGAAA CCTGGCGCAA 2460

TAGATATAGT ACCGCAAGGG AAAGATGAAA AATTATAACC AAGCATAATA TAGCAAGGAC 2520

TAACCCCTAT ACCTTCTGCA TAATGAATTA ACTAGAAATA ACTTTGCAAG GAGAGCCAAA 2580

GCTAAGACCC CCGAAACCAG ACGAGCTACC TAAGAACAGC TAAAAGAGCA CACCCGTCTA 2640

TGTAGCAAAA TAGTGGGAAG ATTTATAGGT AGAGGCGACA AACCTACCGA GCCTGGTGAT 2700

AGCTGGTTGT CCCAGAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2760

AAAAAAAAAA AAAAAAAAAA AAAAAAAA 2788

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 604 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Ser Ser Asn Val Glu Val Phe Ile Pro Val Ser Gln Gly Asn
1 5 10 15

Thr Asn Gly Phe Pro Ala Thr Ala Ser Asn Asp Leu Lys Ala Phe Thr
20 25 30

Glu Gly Ala Val Leu Ser Phe His Asn Ile Cys Tyr Arg Val Lys Leu
35 40 45

```
Lys Ser Gly Phe Leu Pro Cys Arg Lys Pro Val Glu Lys Glu Ile Leu
    50                  55                  60

Ser Asn Ile Asn Gly Ile Met Lys Pro Gly Leu Asn Ala Ile Leu Gly
65                  70                  75                  80

Pro Thr Gly Gly Xaa Lys Ser Ser Leu Leu Asp Val Leu Ala Ala Arg
                85                  90                  95

Lys Asp Pro Ser Gly Leu Ser Gly Asp Val Leu Ile Asn Gly Ala Pro
            100                 105                 110

Arg Pro Ala Asn Phe Lys Cys Asn Ser Gly Tyr Val Val Gln Val Gly
        115                 120                 125

Thr Gln Phe Ile Arg Gly Val Ser Gly Gly Glu Arg Lys Arg Thr Ser
    130                 135                 140

Ile Gly Met Glu Leu Ile Thr Asp Pro Ser Ile Leu Phe Leu Asp Glu
145                 150                 155                 160

Pro Thr Thr Gly Leu Asp Ser Ser Thr Ala Asn Ala Val Leu Leu Leu
                165                 170                 175

Leu Lys Arg Met Ser Lys Gln Gly Arg Thr Ile Ile Phe Ser Ile His
            180                 185                 190

Gln Pro Arg Tyr Ser Ile Phe Lys Leu Phe Asp Ser Leu Thr Leu Leu
        195                 200                 205

Ala Ser Gly Arg Leu Met Phe His Gly Pro Ala Gln Glu Ala Leu Gly
    210                 215                 220

Tyr Phe Glu Ser Ala Gly Tyr His Cys Glu Ala Tyr Asn Asn Pro Ala
225                 230                 235                 240

Asp Phe Phe Leu Asp Ile Ile Asn Gly Asp Ser Thr Ala Val Ala Leu
                245                 250                 255

Asn Arg Glu Glu Asp Phe Lys Ala Thr Glu Ile Ile Glu Pro Ser Lys
            260                 265                 270

Gln Asp Lys Pro Leu Ile Glu Lys Leu Ala Glu Ile Tyr Val Asn Ser
        275                 280                 285

Ser Phe Tyr Lys Glu Thr Lys Ala Glu Leu His Gln Leu Ser Gly Gly
    290                 295                 300

Glu Lys Lys Lys Lys Ile Thr Val Phe Lys Glu Ile Ser Tyr Thr Thr
305                 310                 315                 320

Ser Phe Cys His Gln Leu Arg Trp Val Ser Lys Arg Ser Phe Lys Asn
                325                 330                 335

Leu Leu Gly Asn Pro Gln Ala Ser Ile Ala Gln Ile Ile Val Thr Val
```

340 345 350

Val Leu Gly Leu Val Ile Gly Ala Ile Tyr Phe Gly Leu Lys Asn Asp
355 360 365

Ser Thr Gly Ile Gln Asn Arg Ala Gly Val Leu Phe Phe Leu Thr Thr
370 375 380

Asn Gln Cys Phe Ser Ser Val Ser Ala Val Glu Leu Phe Val Val Glu
385 390 395 400

Lys Lys Leu Phe Ile His Glu Tyr Ile Ser Gly Tyr Tyr Arg Val Ser
405 410 415

Ser Tyr Phe Leu Gly Lys Leu Leu Ser Asp Leu Leu Pro Met Arg Met
420 425 430

Leu Pro Ser Ile Ile Phe Thr Cys Ile Val Tyr Phe Met Leu Gly Leu
435 440 445

Lys Pro Lys Ala Asp Ala Phe Phe Val Met Met Phe Thr Leu Met Met
450 455 460

Val Ala Tyr Ser Ala Ser Ser Met Ala Leu Ala Ile Ala Ala Gly Gln
465 470 475 480

Ser Val Val Ser Val Ala Thr Leu Leu Met Thr Ile Cys Phe Val Phe
485 490 495

Met Met Ile Phe Ser Gly Leu Leu Val Asn Leu Thr Thr Ile Ala Ser
500 505 510

Trp Leu Ser Trp Leu Gln Tyr Phe Ser Ile Pro Arg Tyr Gly Phe Thr
515 520 525

Ala Leu Gln His Asn Glu Phe Leu Gly Gln Asn Phe Cys Pro Gly Leu
530 535 540

Asn Ala Thr Gly Asn Asn Pro Cys Asn Tyr Ala Thr Cys Thr Gly Glu
545 550 555 560

Glu Tyr Leu Val Lys Gln Gly Ile Asp Leu Ser Pro Trp Gly Leu Trp
565 570 575

Lys Asn His Val Ala Leu Ala Cys Met Ile Val Ile Phe Leu Thr Ile
580 585 590

Ala Tyr Leu Lys Leu Leu Phe Leu Lys Lys Tyr Ser
595 600

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2930 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGACTTCCTC GGCTGCGCGG CGCTCGCGCG GAGCTCCCCG GCCGGCGGTG CGTCCCCACG 60

GTCACCATGA AAGACGACTT CGCAGAGGAG GAGGAGGTGC AATCCTTCGG TTACAAGCGG 120

TTTGGTATTC AGGAAGGAAC ACAATGTACC AAATGTAAAA ATAACTGGGC ACTGAAGTTT 180

TCTATCATAT TATTATACAT TTTGTGTGCC TTGCTAACAA TCACAGTAGC CATTTTGGGA 240

TATAAAGTTG TAGAGAAAAT GGACAATGTC ACAGGTGGCA TGGAAACATC TCGCCAAACC 300

TATGATGACA AGCTCACAGC AGTGGAAAGT GACCTGAAAA AATTAGGTGA CCAAACTGGG 360

AAGAAAGCTA TCAGCACCAA CTCAGAACTC TCCACCTTCA GATCAGACAT TCTAGATCTC 420

CGTCAGCAAC TTCGTGAGAT TACAGAAAAA ACCAGCAAGA ACAAGGATAC GCTGGAGAAG 480

TTACAGGCGA GCGGGGATGC TCTGGTGGAC AGGCAGAGTC AATTGAAAGA AACTTTGGAG 540

AATAACTCTT TCCTCATCAC CACTGTAAAC AAAACCCTCC AGGCGTATAA TGGCTATGTC 600

ACGAATCTGC AGCAAGATAC CAGCGTGCTC CAGGGCAATC TGCAGAACCA AATGTATTCT 660

CATAATGTGG TCATCATGAA CTCAACAACC TGAACCTGAC CCAGGTGCAG CAGAGGAACC 720

TCATCACGAA TCTGCAGCGG TCTGTGGATG ACACAAGCCA GGCTATCCAG CGAATCAAGA 780

ACGACTTTCA AAATCTGCAG CAGGTTTTTC TTCAAGCCAA GAAGGACACG GATTGGCTGA 840

AGGAGAAAGT GCAGAGCTTG CAGACGCTGG CTGCCAACAA CTCTGCGTTG GCCAAAGCCA 900

ACAACGACAC CCTGGAGGAT ATGAACAGCC AGCTCAACTC ATTCACAGGT CAGATGGAGA 960

ACATCACCAC TATCTCTCAA GCCAACGAGC AGAACCTGAA AGACCTGCAG GACTTACACA 1020

AAGATGCAGA GAATAGAACA GCCATCAAGT TCAACCAACT GGAGGAACGC TTCCAGCTCT 1080

TTGAGACGGA TATTGTGAAC ATCATTAGCA ATATCAGTTA CACAGCCCAC CACCTGCGGA 1140

CGCTGACCAG CAATCTAAAT GAAGTCAGGA CCACTTGCAC AGATACCCTT ACCAAACACA 1200

CAGATGATCT GACCTCCTTG AATAATACCC TGGCCAACAT CCGTTTGGAT TCTGTTTCTC 1260

TCAGGATGCA ACAAGATTTG ATGAGGTCGA GGTTAGACAC TGAAGTAGCC AACTTATCAG 1320

TGATTATGGA AGAAATGAAG CTAGTAGACT CCAAGCATGG TCAGCTCATC AAGAATTTTA 1380

CAATACTACA AGGTCCACCG GGCCCCAGGG GTCCAAGAGG TGACAGAGGA TCCCAGGGAC 1440

CCCCTGGCCC AACTGGCAAC AAGGGACAGA AAGGAGAGAA GGGGGAGCCT GGACCACCTG 1500

GCCCTGCGGG TGAGAGAGGC CCAATTGGAC CAGCTGGTCC CCCCGGAGAG CGTGGCGGCA 1560

AAGGATCTAA AGGCTCCCAG GGCCCCAAAG GCTCCCGTGG TTCCCCTGGG AAGCCCGGCC 1620

CTCAGGGCCC CAGTGGGGAC CCAGGCCCCC CGGGCCCACC AGGCAAAGAG GGACTCCCCG 1680

GCCCTCAGGG CCCTCCTGGC TTCCAGGGAC TTCAGGGCAC CGTTGGGGAG CCTGGGGTGC 1740

CTGGACCTCG GGGACTGCCA GGCTTGCCTG GGGTACCAGG CATGCCAGGC CCCAAGGGCC 1800

CCCCCGGCCC TCCTGGCCCA TCAGGAGCGG TGGTGCCCCT GGCCCTGCAG AATGAGCCAA 1860

CCCCGGCACC GGAGGACAAT AGCTGCCCGC CTCACTGGAA GAACTTCACA GACAAATGCT 1920

ACTATTTTTC AGTTGAGAAA GAAATTTTTG AGGATGCAAA GCTTTTCTGT GAAGACAAGT 1980

CTTCACATCT TGTTTTCATA AACACTAGAG AGGAACAGCA ATGGATAAAA AAACAGATGG 2040

TAGGGAGAGA GAGCCACTGG ATCGGCCTCA CAGACTCAGA GCGTGAAAAT GAATGGAAGT 2100

GGCTGGATGG GACATCTCCA GACTACAAAA ATTGGAAAGC TGGACAGCCG GATAACTGGG 2160

GTCATGGCCA TGGGCCAGGA GAAGACTGTG CTGGGTTGAT TTATGCTGGG CAGTGGAACG 2220

ATTTCCAATG TGAAGACGTC AATAACTTCA TTTGCGAAAA AGACAGGGAG ACAGTACTGT 2280

CATCTGCATT ATAACGGACT GTGATGGGAT CACATGAGCA AATTTTCAGC TCTCAAAGGC 2340

AAAGGACACT CCTTTCTAAT TGCATCACCT TCTCATCAGA TTGAAAAAAA AAAAGCACTG 2400

AAAGCCAATT ACTGAAAAAA AATTGACAGC TAGTGTTTTT TACCATCCGT CATTACCCAA 2460

AGACTTGGGA ACTAAAATGT TCCCCAGGGT GATATGCTGA TTTTCATTGT GCACATGGAC 2520

TGAATCACAT AGATTCTCCT CCGTCAGTAA CCGTGCGATT ATACAAATTA TGTCTTCCAA 2580

AGTATGGAAC ACTCCAATCA GAAAAAGGTT ATCATTGGTC GTTGAGTTAT GGGAAGAACT 2640

TAAGCATATA CTGTGTAAAC AGTGCCATAC ATTTCTAAAA TCCCAAGTGT AGGAAAAATA 2700

TGCAGACATA CAGATATATA GGCCAACTAT TAGTAATAAT ATGAAATATA CTTAAAGAGC 2760

TTTTAAAACT TTGTATTTTT GTACAAAATA TTTGTCTTTT ACAATTTTTT TCCTTTTTTT 2820

TTTTTTGTCA TTTTACCGAC ATAATACATG GAGCCAAAGA AAACAATAAT GGTACTAATA 2880

AAAACTCCTA GGGTTTCCTG TCAGATTTAA TTCTAAAAAA AAAAAAAAAA 2930

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 208 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Lys Asp Asp Phe Ala Glu Glu Glu Glu Val Gln Ser Phe Gly Tyr
1               5                  10                  15

Lys Arg Phe Gly Ile Gln Glu Gly Thr Gln Cys Thr Lys Cys Lys Asn
            20                  25                  30

Asn Trp Ala Leu Lys Phe Ser Ile Ile Leu Leu Tyr Ile Leu Cys Ala
        35                  40                  45

Leu Leu Thr Ile Thr Val Ala Ile Leu Gly Tyr Lys Val Val Glu Lys
    50                  55                  60

Met Asp Asn Val Thr Gly Gly Met Glu Thr Ser Arg Gln Thr Tyr Asp
65                  70                  75                  80

Asp Lys Leu Thr Ala Val Glu Ser Asp Leu Lys Lys Leu Gly Asp Gln
                85                  90                  95

Thr Gly Lys Lys Ala Ile Ser Thr Asn Ser Glu Leu Ser Thr Phe Arg
            100                 105                 110

Ser Asp Ile Leu Asp Leu Arg Gln Gln Leu Arg Glu Ile Thr Glu Lys
        115                 120                 125

Thr Ser Lys Asn Lys Asp Thr Leu Glu Lys Leu Gln Ala Ser Gly Asp
    130                 135                 140

Ala Leu Val Asp Arg Gln Ser Gln Leu Lys Glu Thr Leu Glu Asn Asn
145                 150                 155                 160

Ser Phe Leu Ile Thr Thr Val Asn Lys Thr Leu Gln Ala Tyr Asn Gly
                165                 170                 175

Tyr Val Thr Asn Leu Gln Gln Asp Thr Ser Val Leu Gln Gly Asn Leu
            180                 185                 190

Gln Asn Gln Met Tyr Ser His Asn Val Val Ile Met Asn Ser Thr Thr
        195                 200                 205
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1589 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCTATATATT TTTTCTAGGA AGGGGTGTTT TTCTTTCTGA TTTAATTCCC TACATTTTTC 60
TCTTTCATAT GAAGTTGCAG ATAATGTTTT TCCTTCGGAT TTTTATTCTT TAAGATTTTT 120
AACCTGTGCA AGACTTTTTC AATGATACAA GTCAAGGAGG ATGAAGATCT TTTTCCACTT 180
CAGTCTTCAC TTTGCTCCAG CTATTGCTAA GAAAGGCACA AACAATGACA GCATATTTAA 240
GGAAGAACCT GGCCGGCTTG GGTCACCGCT GCTGTCTTTC TTGGTTTTGC GTCTACCTGG 300
GAGAGCCCAG CTTTTAGGTT CCCATTGAGG GAAGCATGAG AGAGGATTGT TTGGGGGATG 360
CTGCCAGAGC TTCCAGCTGA CAGTCTCTGC AGAGCGGCTG CCAAGTGGCC TGGTGGCCGT 420
ATGTTGGCAG TTTTTGATGA ATTGGGATTA GGGAATGTTT GTTTACTTGA TAACCGAGTG 480
TCTACAAGGA GAGGTGGCAG CGTGAGGGAA TAGTGCCACC ATAATGAGGA CACAGCCAGC 540
CATCTCTTCC CTGCCACAGA ACCCCAGGCA GTCCCCTTCA GGCTACAGTT TTCCATCTGG 600
ACCGAGGGAC TGGCCGGTGC AGCAGGAGGA GCCGATCACC CTCTGTGGGA ACGAGGATGC 660
CCAGAAGTTC CAGTTACTGT GGCTCCATGG TCCCCTTCTC GATGCGCATC TTGCACGCGG 720
AGCTTCAGCA GTACCTGGGG AACCCACAGG AGTCGCTGGA TAGACTGCAC AAGGTGAAGA 780
CTGTCTGCAG CAAGATCCTG GCCAATTTGG AGCAAGGCTT AGCAGAAGAC GGCGGCATGA 840
GCAGCGTGAC TCAGGAGGGC AGACAAGCCT CTATCCGGCT GTGGAGGTCA CGTCTGGGCC 900
GGGTGATGTA CTCCATGGCA AACTGTCTGC TCCTGATGAA GGATTATGTG CTGGCCGTGG 960
AGGCGTATCA TTCGGTTATC AAGTATTACC CAGAGCAAGA GCCCCAGCTG CTCAGCGGCA 1020
TCGGCCGGAT TTCCCTGCAG ATTGGAGACA TAAAAACAGC TGAAAAGTAT TTTCAAGACG 1080
TTGAGAAAGT AACACAGAAA TTAGATGGAC TACAGGGTAA AATCATGGTT TTGATGAACA 1140
GCGCGTTCCT TCACCTCGGG CAGAATAACT TTGCAGAAGC CCACAGGTTC TTCACAGAGA 1200
TCTTAAGGAT GGATCCAAGA AACGCAGTGG CCAACAACAA CGCTGCCGTG TGTCTGCTCT 1260
ACCTGGGCAA GCTCAAGGAC TCCCTGCGGC AGCTGGAGGC CATGGTCCAG CAGGACCCCA 1320

GGCACTACCT GCACGAGAGC GTGCTCTTCA ACCTGACCAC CATGTACGAG CTGGAGTCCT 1380

CACGGAGCAT GCAGAAGAAA CAGGCCCTGC TGGAGGCTGT CGCCGGCAAG GAGGGGGACA 1440

GCTTCAACAC ACAGTGCCTC AAGCTGGCCT AGCTGCCTCC AACACACTAC GTCAGAAGGA 1500

CCCGGGTCTT TGAAACTGTG TCTTGAAGCT AATGTATTAA TGTGACATGG AGGAACTCAA 1560

TAAAACTCCT GCTTCAAAAA AAAAAAAAA 1589

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Pro Arg Ser Ser Ser Tyr Cys Gly Ser Met Val Pro Phe Ser Met
1 5 10 15

Arg Ile Leu His Ala Glu Leu Gln Gln Tyr Leu Gly Asn Pro Gln Glu
20 25 30

Ser Leu Asp Arg Leu His Lys Val Lys Thr Val Cys Ser Lys Ile Leu
35 40 45

Ala Asn Leu Glu Gln Gly Leu Ala Glu Asp Gly Gly Met Ser Ser Val
50 55 60

Thr Gln Glu Gly Arg Gln Ala Ser Ile Arg Leu Trp Arg Ser Arg Leu
65 70 75 80

Gly Arg Val Met Tyr Ser Met Ala Asn Cys Leu Leu Leu Met Lys Asp
85 90 95

Tyr Val Leu Ala Val Glu Ala Tyr His Ser Val Ile Lys Tyr Tyr Pro
100 105 110

Glu Gln Glu Pro Gln Leu Leu Ser Gly Ile Gly Arg Ile Ser Leu Gln
115 120 125

Ile Gly Asp Ile Lys Thr Ala Glu Lys Tyr Phe Gln Asp Val Glu Lys
130 135 140

Val Thr Gln Lys Leu Asp Gly Leu Gln Gly Lys Ile Met Val Leu Met
145 150 155 160

Asn Ser Ala Phe Leu His Leu Gly Gln Asn Asn Phe Ala Glu Ala His
165 170 175

Arg Phe Phe Thr Glu Ile Leu Arg Met Asp Pro Arg Asn Ala Val Ala
180 185 190

Asn Asn Asn Ala Ala Val Cys Leu Leu Tyr Leu Gly Lys Leu Lys Asp
195 200 205

Ser Leu Arg Gln Leu Glu Ala Met Val Gln Gln Asp Pro Arg His Tyr
210 215 220

Leu His Glu Ser Val Leu Phe Asn Leu Thr Thr Met Tyr Glu Leu Glu
225 230 235 240

Ser Ser Arg Ser Met Gln Lys Lys Gln Ala Leu Leu Glu Ala Val Ala
245 250 255

Gly Lys Glu Gly Asp Ser Phe Asn Thr Gln Cys Leu Lys Leu Ala
260 265 270

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1153 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TATAAAGAGT GACTCTCCTA TGAAGGTAAA GGCCACCCCT CTTCAGTTCC AGTGACTGAG 60

ATACATTTTT CCAATCCTGG GGGCAAATAC AGACACAGCA AGTTCCTTCT TCCCTTTGGA 120

AATTTGGCAG CTGCCTTCAC CAGTGAGCAC AAAGCCACAT TTCAAAGGAA ACTGACAAAT 180

TATCCCCAGC TGCCAGAAGA AGAAATCCTC ACTGGACGGC TTCCTGTTTC CTGTGGTTCA 240

TTATCTGATT GGCTGCAGGG ATGAAAGTTT TTAAGTTCAT AGGACTGATG ATCCTCCTCA 300

CCTCTGCGTT TTCAGCCGGT TCAGGACAAA GTCCAATGAC TGTGCTGTGC TCCATAGACT 360

GGTTCATGGT CACAGTGCAC CCCTTCATGC TAAACAACGA TGTGTGTGTA CACTTTCATG 420

AACTACACTT GGGCCTGGGT TGCCCCCCAA ACCATGTTCA GCCACACGCC TACCAGTTCA 480

CCTACCGTGT TACTGAATGT GGCATCAGGG CCAAAGCTGT CTCTCAGGAC ATGGTTATCT 540

ACAGCACTGA GATACACTAC TCTTCTAAGG GCACGCCATC TAAGTTTGTG ATCCCAGTGT 600

CATGTGCTGC CCCCCAAAAG TCCCCATGGC TCACCAAGCC CTGCTCCATG AGAGTAGCCA 660

GCAAGAGCAG GGCCACAGCC CAGAAGGATG AGAAATGCTA CGAGGTGTTC AGCTTGTCAC 720

AGTCCAGTCA AAGGCCCAAC TGCGATTGTC CACCTTGTGT CTTCAGTGAA GAAGAGCATA 780

CCCAGGTCCC TTGTCACCAA GCAGGGGCTC AGGAGGCTCA ACCTCTGCAG CCATCTCACT 840

TTCTTGATAT TTCTGAGGAT TGGTCTCTTC ACACAGATGA TATGATTGGG TCCATGTGAT 900

CCTCAGGTTT GGGGTCTCCT GAAGATGCTA TTTCTAGAAT TAGTATATAG TGTACAAATG 960

TCTGACAAAT AAGTGCTCTT GTGACCCTCA TGTGAGCACT TTTGAGAAAG AGAAACCTAT 1020

AGCAACTTCA TGAATTAAGC CTTTTTCTAT ATTTTTATAT TCATGTGTAA ACAAAAAATA 1080

AAATAAAATT CTGATCGCAT AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1140

AAAAAAAAAA AAA 1153

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 212 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Lys Val Phe Lys Phe Ile Gly Leu Met Ile Leu Leu Thr Ser Ala
1 5 10 15

Phe Ser Ala Gly Ser Gly Gln Ser Pro Met Thr Val Leu Cys Ser Ile
20 25 30

Asp Trp Phe Met Val Thr Val His Pro Phe Met Leu Asn Asn Asp Val
35 40 45

Cys Val His Phe His Glu Leu His Leu Gly Leu Gly Cys Pro Pro Asn
50 55 60

His Val Gln Pro His Ala Tyr Gln Phe Thr Tyr Arg Val Thr Glu Cys
65 70 75 80

Gly Ile Arg Ala Lys Ala Val Ser Gln Asp Met Val Ile Tyr Ser Thr
85 90 95

Glu Ile His Tyr Ser Ser Lys Gly Thr Pro Ser Lys Phe Val Ile Pro
100 105 110

Val Ser Cys Ala Ala Pro Gln Lys Ser Pro Trp Leu Thr Lys Pro Cys
115 120 125

Ser Met Arg Val Ala Ser Lys Ser Arg Ala Thr Ala Gln Lys Asp Glu
130 135 140

Lys Cys Tyr Glu Val Phe Ser Leu Ser Gln Ser Ser Gln Arg Pro Asn
145 150 155 160

Cys Asp Cys Pro Pro Cys Val Phe Ser Glu Glu Glu His Thr Gln Val
165 170 175

Pro Cys His Gln Ala Gly Ala Gln Glu Ala Gln Pro Leu Gln Pro Ser
180 185 190

His Phe Leu Asp Ile Ser Glu Asp Trp Ser Leu His Thr Asp Asp Met
195 200 205

Ile Gly Ser Met
210

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4285 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTTAATCTGT GTCTCCAGCA TTTATTTTTT TGTTTGTGTC ATCGGGTTCC TGGTTTTCTT 60

TTAAGACATA GTCAACTGTG TGGACCTGTA GGTTTGGGGC AGCAACCAAT TCCATTGTTT 120

TCCTTTTTGT CAAATCCAAG AGAAAATATA CCATAAGGAG CTAGAAGATT CTAGTTCACA 180

GCCTTTTGAA TCTTCATGGC CTTTGAATCC TCATGGCCTC TGAAATCTGA ATCAGTTTTC 240

TCCCAGGARG TCTCTGGGGG CTGAGCTGCT ACAGGGGCAR ARGGTGGGGT GGGGTTGGGT 300

GGGARAATCA TCCTGGCACT TCATCGTGCA TGCTATTTCG GGCAGCATCT TTTTTTTTTT 360

ATTTTATTAT TATTTTTTTT CCTGATGCTT GAGTTATGAA TGAGGATGAC CTCTGCAATC 420

ATGATGTCTC CCATAGACTC TGTTCCTTGT TCCTTTGCCA GCTTTCTCAT GCATGGTCCT 480

AACACTTCCA TGATTTAATC TGCTGCAGGA CCATAGTCTT CAGCCACCTC AGCAATAACT 540

TGTTAGAACA TTAAAAGGAA GTAAATTGAG AACAACTTGT TGCCATCCCA TTTTCATTAG 600

AAATCAGACA TCTTAGAGAT GTCAAGAAAG CAGCTAGCAG CTAGGGGGTA TGGGGACCTG 660

TCCTGCTCAC ACTGCTGTGT GTCAGACCAG ACCTGATCCT GGAGCTCAGG ACCCTAGAGA 720

GCCCTGATCT CTGGAACTCT TGCCACGTTG TTGCTGAGGC AGCTGAAGTC CCCATCTCCC 780

ACCATAACAA TCACAAATAG ACAGTAGTGG AGCCAGCATC CCCAGGCCCC TTTTTGTGTA 840

AGCAGAAAGG GAGCTGTGAG CCTTGCCCTG TTTGCAGGTG TCAAGTGCCT CTCCCTGCCT 900

GTACTTCTCC CCTTCCTCTG AGCAGAGCTT TGGTAGCTGT TGCCAATGCA AAGAAATGTA 960

AAGCAGCAAA AGAAGACAGC AGGTTCTGAC CTGAGGAGGG AAACCAAATT TATCCCACAA 1020

AGGCCCATTA ACCCCACCCC CCTCGCCTCC CACCCCCAGA CTGGATCCAC TACTGGCCCA 1080

AGAATACTGA TGAGAAACCT AGTCTGGATT GGGTCGGAAG CTGGAATTTG GTGCTCTGCA 1140

GACCAGTGCT CAAAATTGTG GTTATTTTTG AGGACTCGCC TTCAATCCAG AACATTTGCG 1200

TTTCACCTTC CTCGCCCAGA TCCAGTTAAC AAGGTAGCTC ATCACTTCTT GCATCTGTTG 1260

AGTGACATGC TGGATTTTAA TTTTTATTGT GGTTGTACTT GGATGCAAGG AATATGTTTT 1320

GTTCCTCCCA ATTTAGCGCA CCATCCTGGG AAGTGCATGT CTCAGACCAA CTCCACCTTC 1380

ACCTTCACCA CCTGTCGCAT CCTGCATCCT TCAGATGAGC TCACTCGGGT CACACCAAGC 1440

CTTAACTCAG CCCCAACTCC AGCTTGTGGC AGCACCAGCC ACTTGAAATC CACGCCGGTG 1500

GCCACACCAT GCACTCCACG GAGACTGAGC CTGGCTGAGT CCTTCACTAA CACCCGTGAG 1560

TCCACGACCA CCATGAGCAC ATCCCTGGGG CTCGTGTGGC TGTTGAAGGA GCGGGGCATT 1620

TCTGCTGCCG TGTACGACCC CCAGAGCTGG GACAGGGCCG GCCGGGGCTC CCTCCTGCAC 1680

TCCTACACGC CCAAGATGGC TGTGATCCCC TCTACTCCGC CGAACTCGCC TATGCAGACA 1740

CCCACATCCT CCCCACCCTC CTTTGAGTTC AAGTGCACGA GCCCTCCCTA CGACAATTTC 1800

CTGGCTTCCA AGCCAGCCAG CTCCATCCTG AGGGAAGTGA GAGAAAAGAA CGTCCGCAGC 1860

AGCGAGAGCC AGACCGACGT GTCCGTCTCC AACCTCAACC TCGTGGACAA AGTCAGGAGG 1920

TTTGGGGTGG CCAAAGTGGT GAACTCAGGG CGAGCCCATG TCCCCACCTT GACTGAGGAG 1980

CAGGGACCCC TCCTCTGTGG GCCCCCGGGG CCAGCACCAG CCCTTGTTCC CAGAGGCCTG 2040

GTACCTGAGG GCCTGCCCCT CAGATGCCCC ACTGTCACCA GTGCCATCGG TGGGCTGCAG 2100

CTCAATAGTG GCATCCGGCG GAATCGCAGC TTCCCCACCA TGGTGGGATC TAGCATGCAG 2160

ATGAAAGCTC CTGTGACTCT CACCTCGGGC ATCTTGATGG GTGCTAAGCT CTCCAAACAA 2220

ACTAGCTTAC GGTGAGGACT GGAGGGGGGC CGGTTGCCCT AGAGGAGACC CACGTTCTCT 2280

CTTGCTCCCA CCTCCCTCTC TTCCCCCCAC AGTGCACTCC CTCCCTCTGC CCTTCTCTGT 2340

CCACCCCCTC CTAAGCTAGA CAAATCAACC TTGTGCCTAA TGGAGGAAGT GTGGAAACTT 2400

TGTAAAATGT GTACATAGGA CTTGGAGACC TTGTGTCCGC CCTGCTCTTT CTTCCGATCC 2460

CACAGGAAGT GCCCCTGCAC TGTCATCACT CTCACGAGGA CGTCACCTGT GCTAACCTGG 2520

GGGAAGGTGG GGTCCTTTCT TCTTTCCTTT TGAGAAGCAC TGAAACTCCC AAGTGTGTTC 2580

TTATCCCATG GATAGGAAAC CAGTGAATTC CGTGGCTGGC ACACCACGAG CTGTCATGCG 2640

GCACGGGTCA TAACACATCT GGGTGTCATC GGACACCTCA CCTCGCCCAC CCTGTAGGAG 2700

CGTAAGGAGC CTCCATCCTC AGCCACGTGC AGCTGACGTG GCTTTCCTGA TCGGAGGGCT 2760

TTTCTTTTAT GGGTGGCCCA GCTTCTTCAA GACCTTCACT GCTCTGCCTC AGTGGACAGT 2820

CGTTTCTTTT TTGAGGTGTG ACCTTTTGTT TTCATGCCTT CCCCTTGAAG TCATCCTGTG 2880

TTTTGTAATC AGCTGTCAGG CCAAATGTCT GACCCGAAAG AGAATGTATT TACACTCATG 2940

CTGCGTTGTT CAGCAGCCCC TCTGTGTTCT GTGTGATTTG TTTTATTTTT CCTTTTTTTT 3000

ACATATATAT GCAGGGAAGT AATGGTACTG GTAGTGTATG TTTTCTATGT GGTTCAAATA 3060

TGAATTTCGA ACACACCAAG CCGCTAATGA GATAGCAGCT TTTTTCTGGG ACCCAGAGTC 3120

ACAACCAAAT TGATTTAAGA CCGGACCCAA GACACCTTTA ACAATAGGAC TGAAAGGAAA 3180

AAGGATAGGG AAAAAGCTTA TTAAAGAAAT GTGTCAACAC CAAATGTAGA GGGGAAGAAC 3240

CACAACCAGG CATAATACCA AACCGGTTCC AGGGGGAAAC AAGGCTTTGG TATTCCGCTG 3300

GCTCCAGCGC TTTTTCTGAA ACCCGAGGCT GGCCAGGGTG CTGTCACCGT GTGGTCTTTG 3360

ATTGCAGCCA TTCAATGCCC ACATGCTTTT CCTTCTTGTT TCAGAACAGC ACATGGTCAC 3420

AACAAGATAT TTTCTTTCCC TCCAAAGCCT TTTGTCTCCT TGTGCCTCTT TTTATCCTTA 3480

GGAAAAGATC CAGGTGCTTG TGAAAAGAAT CATGAATGCA ACAAGGGAGG CTGGTCCTGT 3540

TGCTGTCGCC GATTAAGTTT TAAACTTTTA TTTATTATTT ATGTCTGCCG TATTTTAAAT 3600

AAACATTCTC GTTCCTTCCA GTTCCAGTCA TAGTGTGTCT GTGGCATTCC AGTCCAACCA 3660

TGTGACTTAT TTATTCTAAT TTGAGGGCTG CACTGTACAC CATGGTGTCC TGTGACACCG 3720

TGTTCCAGAC ATTTATGGAA GGAAAACATC CCATATAAAT GAAACTGTCA TGCTGTGTCC 3780

TCCCCGGCAG CAGAAGATGT GTCCTTCCAT TGAGTGAGGG TAACCTTATG TCCACCAAGG 3840

ATACTTTGAG AAAGCCCCTA AGGAACAAGC CTCAGTCCCA CGGTTTCAGA CTATTTATTC 3900

TCTGAACACA AGAGTATTGG TTAATTATGT TCTCAGCTCT CCCTGCTGTT GTATGTGTGC 3960

ATTCACTGCA AGTAACTTAT ATCTTTTTAT TTGAATGTAT TTTAAAGCAG TAGATAGAAT 4020

AACAAAGGAA TATGAAAACC ATGGACTGAA TGGACCATTT TATGTATTCA GAGAGAGAAG 4080

CCACTCATCA TTGCCAGAAA TACCATGTAA AAATTGGCAG TTCAGAGGTT GCAATACTTA 4140

GTATAGTAAA TAAATAAACG GTCAACATTG TGCAACCACT ACCAAAAAGT GTGTTGTAAT 4200

GCATCAAAAA TCAACACAAT TTTATTCACT AATGAGTATC AATAAAATAA GTTCAAATGA 4260

TGGAAACCAC AAAAAAAAAA AAAAA 4285

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 429 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Gln Arg Asn Val Lys Gln Gln Lys Lys Thr Ala Gly Ser Asp Leu
1 5 10 15

Arg Arg Glu Thr Lys Phe Ile Pro Gln Arg Pro Ile Asn Pro Thr Pro
20 25 30

Leu Ala Ser His Pro Gln Thr Gly Ser Thr Thr Gly Pro Arg Ile Leu
35 40 45

Met Arg Asn Leu Val Trp Ile Gly Ser Glu Ala Gly Ile Trp Cys Ser
50 55 60

Ala Asp Gln Cys Ser Lys Leu Trp Leu Phe Leu Arg Thr Arg Leu Gln
65 70 75 80

Ser Arg Thr Phe Ala Phe His Leu Pro Arg Pro Asp Pro Val Asn Lys
85 90 95

Val Ala His His Phe Leu His Leu Leu Ser Asp Met Leu Asp Phe Asn
100 105 110

Phe Tyr Cys Gly Cys Thr Trp Met Gln Gly Ile Cys Phe Val Pro Pro
115 120 125

Asn Leu Ala His His Pro Gly Lys Cys Met Ser Gln Thr Asn Ser Thr
130 135 140

Phe Thr Phe Thr Thr Cys Arg Ile Leu His Pro Ser Asp Glu Leu Thr

```
145                 150                 155                 160

Arg Val Thr Pro Ser Leu Asn Ser Ala Pro Thr Pro Ala Cys Gly Ser
                165                 170                 175

Thr Ser His Leu Lys Ser Thr Pro Val Ala Thr Pro Cys Thr Pro Arg
            180                 185                 190

Arg Leu Ser Leu Ala Glu Ser Phe Thr Asn Thr Arg Glu Ser Thr Thr
        195                 200                 205

Thr Met Ser Thr Ser Leu Gly Leu Val Trp Leu Leu Lys Glu Arg Gly
    210                 215                 220

Ile Ser Ala Ala Val Tyr Asp Pro Gln Ser Trp Asp Arg Ala Gly Arg
225                 230                 235                 240

Gly Ser Leu Leu His Ser Tyr Thr Pro Lys Met Ala Val Ile Pro Ser
                245                 250                 255

Thr Pro Pro Asn Ser Pro Met Gln Thr Pro Thr Ser Ser Pro Pro Ser
            260                 265                 270

Phe Glu Phe Lys Cys Thr Ser Pro Pro Tyr Asp Asn Phe Leu Ala Ser
        275                 280                 285

Lys Pro Ala Ser Ser Ile Leu Arg Glu Val Arg Glu Lys Asn Val Arg
    290                 295                 300

Ser Ser Glu Ser Gln Thr Asp Val Ser Val Ser Asn Leu Asn Leu Val
305                 310                 315                 320

Asp Lys Val Arg Arg Phe Gly Val Ala Lys Val Val Asn Ser Gly Arg
                325                 330                 335

Ala His Val Pro Thr Leu Thr Glu Glu Gln Gly Pro Leu Leu Cys Gly
            340                 345                 350

Pro Pro Gly Pro Ala Pro Ala Leu Val Pro Arg Gly Leu Val Pro Glu
        355                 360                 365

Gly Leu Pro Leu Arg Cys Pro Thr Val Thr Ser Ala Ile Gly Gly Leu
    370                 375                 380

Gln Leu Asn Ser Gly Ile Arg Arg Asn Arg Ser Phe Pro Thr Met Val
385                 390                 395                 400

Gly Ser Ser Met Gln Met Lys Ala Pro Val Thr Leu Thr Ser Gly Ile
                405                 410                 415

Leu Met Gly Ala Lys Leu Ser Lys Gln Thr Ser Leu Arg
            420                 425
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3751 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ACTTTGAATT TTTTATTTGT GAAATTAAAA ATATGGTATT ATATATATAT AAACTTCTAT 60
TCCTCTATAA ATATAGATGA TTTTGTGATA GTGAACAGAA TAAATGTATA CCAAATTCAA 120
AGACCAATAT CATTTTAGCG TATGACAGAC ATAGATAAAT TTAGGTCCTA AGTACCGGCA 180
TTTTGATAAA TTCTTAAAGT TTAAAACAAT ACAATCAGGA GGATTGCTTT TCTCCTCTTC 240
TTCACAGAGA ACTAAAGTGA ATATTTTTAA ATGGCTTTGA AAGATTTACA TTTGACACAT 300
TTCTGTAAAT CCAAAAGAGG AGCACACAGG GATTTAATGC AGTAGACCTG CACACATTTT 360
CCCTTTAGCA TGCATGCCCA TATTTTGTTT ATTTCAGGCG CTATCTCCCC GTCAATTATT 420
CCACCTTCTT TACCTCCTGA AATCTTACCA GGTTATTATT GGTGGTGTGA ATTGTTCCCC 480
CCTCAGAATG TGCTGCTGAA TAATAATCGT AATAAAATGT TGAAAGTGTA CAACTTTTAC 540
ATTTTAAAGT TTCTGATATA TGTCTAGTTA TTTGATTAAA AATAAGAAAA TAGCACTTCA 600
TTTTGAGGAA GTCCATGACA CTGAAATATC CTTCAAGTTT TCAATTTCTG TTTACGTTTT 660
GCTGTCTTGT TAAGGAAAGC AAACATCAAC TCCTTAACAA AGCTTTCCAG GTGACCTCAA 720
CATTTCCATT TTACAGACCG GTAAAATCTA AGCGCAGGCT GTCTCATTCT CAAAGGCAAG 780
GTTGCCAGGC ATCCGTATGC AATTAGAATT AACATTTTAT AACCCATATC TTCAGTCTCT 840
TCCAACCCAC ACAAAGCTTC ATGCTTCTTC CCAAATCTCA GTAACCACAT CTTTCCATGA 900
CGCTGGCCAA ACCCATACCA GGTTTTAGAC ACTAGAGAAT GAAATGAGCT CACCCCTCAA 960
AAATTAGACT TCAAAAAGTT TGGCATTGGT TATCTCACTC ACCCTGTAAC CAACTAAGGT 1020
GGGAGAAGGG AGTGTCTGGC GTTGAAGGTG ACCGTGGAGG GAGGCTGAGA CTGCCAGCGC 1080
CCACACCCGT GGGCCCCCAT GAAGTTGGAG GAAAGTTCTG GACAGTTAAA AATCCAGCTT 1140
CAGGAAGTCG AAGGGACGGG CCTTCGCAAT CCACCGCCGA GCAAGGGAGG AATTGTAATG 1200
TATGGGGGCC CTCCTCCAGA TTTGGAAGGT TTGTGGAGTT CTGTACCTTA AGAGCCCCTA 1260

CCTCAAGCCA GGAAAGAAAG GGAGGGGACA GAAGGAGGGG GAGGGGGCAA AAGGAGGAGG 1320

CGGGAAGTGA CCCTGGCAGC GCAGCCCTAG TCGCACCCCG CAGTGCTGAA CTCGCCCCGG 1380

AGCTGGCGCC CAGCCGTCCC GAGCACCCGT GGTAGGGAGA GGCGCGCGAG GACGACCAGG 1440

AGCGCTGTGC GGTTGCACAC CAGTTTTAGC TCCTTTGCAA TACTCCGAAA AGGGCAAGAA 1500

GAAAAGCCTC AAATGGTTAA ACCGCCCTAA ATAATTAAAA ACTTTTGAAA AAGAAAAACG 1560

CGTGATCGGT CGTCATTTAA ATACAAATAT ACTTACAAAA ATCCTACACA GGCTATTTAC 1620

AATCATAAAA GCGAACAGTC CTGGTACCAG AGTGTGAGGG CAAGAGGTCT GTCCATCCTC 1680

CCTCTGGCAG TCGGGCCCTC GTGTCCTTTT GCCTCAGGGA CGGAAGCTTT TGCAGGAGCT 1740

GAGTTGTTCT AGGCCTCTTT GGCCGAATTC GGCCAAAGAG GCCTAATTCC TTCCTCGGTT 1800

ATTTCATTCA GAGAATATTT ATGAAATGCC TACTGTGTGC AAGTCATCCA TCCTTGAAAA 1860

GGCCACTTCT CAGTGAGGGA GAGATGTAGT GGATTCTGTG AGACATACCT GCTGGAGTTG 1920

AAGCAGTAAA TAGCATGTCT TTCCCCTCCC CGATCTTAAG GTGTGTTTTC TAGAAAAGTT 1980

CCCTAATGGA ATTCATGAGT TTGGGGGTCT CAGTCACCCG CTTGCCTGTA GGATTCCATT 2040

TGATGATTCT GGATTTTTGC TGTTTGTTAT TGCCCTTAGA GGGGCTCTGA GTATCTACTT 2100

GTGGGTGGCC ATTTCCTGAC ATCTGCATGT ACCTCGTGGA ATTCAGCCAG CTTCATGTTG 2160

CAAATCAGAA AGCTGACCCC AAGACTGCAA ATCAATGAAG GTATTGGCAT TGTTAAGGTC 2220

GTAGCGTAGA CAACAGCAGT CATAAATAAT TAGGCAGGAA CTTAACCCAA ATCTAGTTCT 2280

TTGACCACCT CTACCACCAG AACCCAGCAG ACACTCACAT CTCCTGATAA GAGTTGCTGG 2340

ACTCGATGTT TTTGTTTTGC ATTTTCTCCT CTCCTTCCCC ACTTACTCAG AGAATTTAAA 2400

GTCTGTAGAG TCAGCACAGC CCCATCAGTC CAGGAACTTC CCACCACCAG CCCTTGACTG 2460

TCCCATTAAC TGACATGGTC AGATTTCCAG CTCCCCCTAC TCCCTGCTGT GAAACAATCC 2520

CTCTCCYTGT GAGAGGAAAY TGCGCGSGAA GGYTAAGGGA GTGTGGCGGG CGGYTCCGGG 2580

AGCCAACATG CCTCGGTATG CGCAGCTGKT CATGGSCCCC GCGGGCAGCG GGAAGAGCAC 2640

YTACTGTGCC ACCATGGTCC AGCACTGTGA AGCCYTCAAC CGGTCTGTCC AAGTTGTAAA 2700

CCTGGATCCA GCAGCAGAAC ACTTCAAYTA CTCCGTGATG GCTGACATCC GGGAACTGAT 2760

CGAGGTGGAT GATGTAATGG AGGATGATTY TYTGCGATTC GGTCCCAACG GAGGATTGGT 2820

ATTTTGCATG GAGTACTTTG CCAATAATTT TGACTGGCTG GAGAACTGTC TTGGCCATGT 2880

AGAGGACGAC TATATCCTTT TTGATTGTCC AGGTCAGATT GAGTTGTACA CTCACCTGCC 2940

TGTGATGAAA CAGCTGGTCC AGCAGCTCGA GCAGTGGGAG TTCCGAGTCT GTGGAKTTTY 3000

TYTTGTTGAT TCTCAGTTCA TGGTGGAGTC ATTCAAGTTT ATTTCTGGCA TCTTGGCAGC 3060

CCTGAGTGCC ATGATCTCTC TAGAAATTCC GCAAGTCAAC ATCATGACAA AAATGGATCT 3120

GCTGAGTAAA AAAGCAAAAA AGGAAATTGA GAAATTTTTA GATCCAGACA TGTATTCTTT 3180

ATTAGAAGAT TCTACAAGTG ACTTAAGAAG CAAAAAATTC AAGAAACTGA CTAAAGCTAT 3240

ATGTGGACTG ATTGATGACT ACAGCATGGT TCGATTTTTA CCTTACGATC AGTCAGATGA 3300

AGAAAGCATG AACATTGTAT TGCAGCATAT TGATTTTGCC ATTCAATATG GAGAAGACCT 3360

AGAATTTAAA GAACCAAAGG AACGTGAAGA TGAGTCTTCC TCTATGTTTG ACGAATATTT 3420

TCAAGAATGC CAGGATGAAT GAAGAGTTTA CTAAAAGTAA CCATCTAAAG AGCTTGTGGC 3480

CAAACCAGCA GAACATTCTT CTYTTCAAAG GATGCAATAG TAGAAAGCTA CTTATTTTAA 3540

TGAAAAAAAG TAAAACTTCG TTCTTTATCA GCCTCATGCC TGAATCAAAT TTTTAATTAT 3600

TCTGAAACTG CTGCTGTTTA AAGTGGAATC TTTTAGTATT ATAACAGCAT CACTTTAGAT 3660

TTTGTAAGTC AAAATTGAAA TGAATGCACA TAGATTTATA TATAAATTAG CACCTGAGCT 3720

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A 3751

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 284 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Pro Arg Tyr Ala Gln Leu Xaa Met Xaa Pro Ala Gly Ser Gly Lys
1 5 10 15

Ser Thr Tyr Cys Ala Thr Met Val Gln His Cys Glu Ala Xaa Asn Arg
20 25 30

Ser Val Gln Val Val Asn Leu Asp Pro Ala Ala Glu His Phe Asn Tyr
35 40 45

Ser Val Met Ala Asp Ile Arg Glu Leu Ile Glu Val Asp Asp Val Met
50 55 60

```
Glu Asp Asp Xaa Leu Arg Phe Gly Pro Asn Gly Gly Leu Val Phe Cys
65                  70                  75                  80

Met Glu Tyr Phe Ala Asn Asn Phe Asp Trp Leu Glu Asn Cys Leu Gly
                85                  90                  95

His Val Glu Asp Asp Tyr Ile Leu Phe Asp Cys Pro Gly Gln Ile Glu
            100                 105                 110

Leu Tyr Thr His Leu Pro Val Met Lys Gln Leu Val Gln Gln Leu Glu
        115                 120                 125

Gln Trp Glu Phe Arg Val Cys Gly Xaa Xaa Xaa Val Asp Ser Gln Phe
    130                 135                 140

Met Val Glu Ser Phe Lys Phe Ile Ser Gly Ile Leu Ala Ala Leu Ser
145                 150                 155                 160

Ala Met Ile Ser Leu Glu Ile Pro Gln Val Asn Ile Met Thr Lys Met
                165                 170                 175

Asp Leu Leu Ser Lys Lys Ala Lys Lys Glu Ile Glu Lys Phe Leu Asp
            180                 185                 190

Pro Asp Met Tyr Ser Leu Leu Glu Asp Ser Thr Ser Asp Leu Arg Ser
        195                 200                 205

Lys Lys Phe Lys Lys Leu Thr Lys Ala Ile Cys Gly Leu Ile Asp Asp
    210                 215                 220

Tyr Ser Met Val Arg Phe Leu Pro Tyr Asp Gln Ser Asp Glu Glu Ser
225                 230                 235                 240

Met Asn Ile Val Leu Gln His Ile Asp Phe Ala Ile Gln Tyr Gly Glu
                245                 250                 255

Asp Leu Glu Phe Lys Glu Pro Lys Glu Arg Glu Asp Glu Ser Ser Ser
            260                 265                 270

Met Phe Asp Glu Tyr Phe Gln Glu Cys Gln Asp Glu
        275                 280
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TNCAGGCCTT GCGTTCCTAG CTGCTCTGC 29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GNGCTGTGAG TTTATCCACA AAGGAACAG 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GNATAGGAGG TCCCAAGTTA TCAAGGTTT 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GNTTTCCTGG TTCTTGGTCA GGTTTCCTC 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CNAGATGCAA TGGTTGTGAG ATTGACCAA 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GNCACTTTCC ACTGCTGTGA GCTTGTCAT 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ANCAGACAGT TTGCCATGGA GTACATCAC 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TNATGAACCA CAGGAAACAG GAAGCCGTC 29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TNAAGGTGAA GGTGGAGTTG GTCTGAGAC 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GNCAGAAATA AACTTGAATG ACTCCACCA 29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 457 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Met Asn Ser Gln Leu Asn Ser Phe Thr Gly Gln Met Glu Asn Ile Thr
1               5                  10                  15

Thr Ile Ser Gln Ala Asn Glu Gln Asn Leu Lys Asp Leu Gln Asp Leu
            20                  25                  30

His Lys Asp Ala Glu Asn Arg Thr Ala Ile Lys Phe Asn Gln Leu Glu
        35                  40                  45

Glu Arg Phe Gln Leu Phe Glu Thr Asp Ile Val Asn Ile Ile Ser Asn
    50                  55                  60

Ile Ser Tyr Thr Ala His His Leu Arg Thr Leu Thr Ser Asn Leu Asn
65                  70                  75                  80

Glu Val Arg Thr Thr Cys Thr Asp Thr Leu Thr Lys His Thr Asp Asp
                85                  90                  95

Leu Thr Ser Leu Asn Asn Thr Leu Ala Asn Ile Arg Leu Asp Ser Val
            100                 105                 110

Ser Leu Arg Met Gln Gln Asp Leu Met Arg Ser Arg Leu Asp Thr Glu
        115                 120                 125

Val Ala Asn Leu Ser Val Ile Met Glu Glu Met Lys Leu Val Asp Ser
    130                 135                 140

Lys His Gly Gln Leu Ile Lys Asn Phe Thr Ile Leu Gln Gly Pro Pro
145                 150                 155                 160

Gly Pro Arg Gly Pro Arg Gly Asp Arg Gly Ser Gln Gly Pro Pro Gly
                165                 170                 175

Pro Thr Gly Asn Lys Gly Gln Lys Gly Glu Lys Gly Glu Pro Gly Pro
            180                 185                 190

Pro Gly Pro Ala Gly Glu Arg Gly Pro Ile Gly Pro Ala Gly Pro Pro
        195                 200                 205

Gly Glu Arg Gly Gly Lys Gly Ser Lys Gly Ser Gln Gly Pro Lys Gly
    210                 215                 220
```

```
Ser Arg Gly Ser Pro Gly Lys Pro Gly Pro Gln Gly Pro Ser Gly Asp
225                 230                 235                 240

Pro Gly Pro Pro Gly Pro Pro Gly Lys Glu Gly Leu Pro Gly Pro Gln
                245                 250                 255

Gly Pro Pro Gly Phe Gln Gly Leu Gln Gly Thr Val Gly Glu Pro Gly
            260                 265                 270

Val Pro Gly Pro Arg Gly Leu Pro Gly Leu Pro Gly Val Pro Gly Met
        275                 280                 285

Pro Gly Pro Lys Gly Pro Pro Gly Pro Pro Gly Pro Ser Gly Ala Val
    290                 295                 300

Val Pro Leu Ala Leu Gln Asn Glu Pro Thr Pro Ala Pro Glu Asp Asn
305                 310                 315                 320

Ser Cys Pro Pro His Trp Lys Asn Phe Thr Asp Lys Cys Tyr Tyr Phe
                325                 330                 335

Ser Val Glu Lys Glu Ile Phe Glu Asp Ala Lys Leu Phe Cys Glu Asp
            340                 345                 350

Lys Ser Ser His Leu Val Phe Ile Asn Thr Arg Glu Glu Gln Gln Trp
        355                 360                 365

Ile Lys Lys Gln Met Val Gly Arg Glu Ser His Trp Ile Gly Leu Thr
    370                 375                 380

Asp Ser Glu Arg Glu Asn Glu Trp Lys Trp Leu Asp Gly Thr Ser Pro
385                 390                 395                 400

Asp Tyr Lys Asn Trp Lys Ala Gly Gln Pro Asp Asn Trp Gly His Gly
                405                 410                 415

His Gly Pro Gly Glu Asp Cys Ala Gly Leu Ile Tyr Ala Gly Gln Trp
            420                 425                 430

Asn Asp Phe Gln Cys Glu Asp Val Asn Asn Phe Ile Cys Glu Lys Asp
        435                 440                 445

Arg Glu Thr Val Leu Ser Ser Ala Leu
    450                 455
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 542 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Cys Gly His His Glu Leu Asn Asn Leu Asn Leu Thr Gln Val Gln Gln
1               5                  10                  15

Arg Asn Leu Ile Thr Asn Leu Gln Arg Ser Val Asp Asp Thr Ser Gln
            20                  25                  30

Ala Ile Gln Arg Ile Lys Asn Asp Phe Gln Asn Leu Gln Gln Val Phe
        35                  40                  45

Leu Gln Ala Lys Lys Asp Thr Asp Trp Leu Lys Glu Lys Val Gln Ser
    50                  55                  60

Leu Gln Thr Leu Ala Ala Asn Asn Ser Ala Leu Ala Lys Ala Asn Asn
65                  70                  75                  80

Asp Thr Leu Glu Asp Met Asn Ser Gln Leu Asn Ser Phe Thr Gly Gln
                85                  90                  95

Met Glu Asn Ile Thr Thr Ile Ser Gln Ala Asn Glu Gln Asn Leu Lys
            100                 105                 110

Asp Leu Gln Asp Leu His Lys Asp Ala Glu Asn Arg Thr Ala Ile Lys
        115                 120                 125

Phe Asn Gln Leu Glu Glu Arg Phe Gln Leu Phe Glu Thr Asp Ile Val
    130                 135                 140

Asn Ile Ile Ser Asn Ile Ser Tyr Thr Ala His His Leu Arg Thr Leu
145                 150                 155                 160

Thr Ser Asn Leu Asn Glu Val Arg Thr Thr Cys Thr Asp Thr Leu Thr
                165                 170                 175

Lys His Thr Asp Asp Leu Thr Ser Leu Asn Asn Thr Leu Ala Asn Ile
            180                 185                 190

Arg Leu Asp Ser Val Ser Leu Arg Met Gln Gln Asp Leu Met Arg Ser
        195                 200                 205

Arg Leu Asp Thr Glu Val Ala Asn Leu Ser Val Ile Met Glu Glu Met
    210                 215                 220

Lys Leu Val Asp Ser Lys His Gly Gln Leu Ile Lys Asn Phe Thr Ile
225                 230                 235                 240

Leu Gln Gly Pro Pro Gly Pro Arg Gly Pro Arg Gly Asp Arg Gly Ser
                245                 250                 255

Gln Gly Pro Pro Gly Pro Thr Gly Asn Lys Gly Gln Lys Gly Glu Lys
            260                 265                 270
```

```
Gly Glu Pro Gly Pro Pro Gly Pro Ala Gly Glu Arg Gly Pro Ile Gly
        275                 280                 285

Pro Ala Gly Pro Pro Gly Glu Arg Gly Gly Lys Gly Ser Lys Gly Ser
    290                 295                 300

Gln Gly Pro Lys Gly Ser Arg Gly Ser Pro Gly Lys Pro Gly Pro Gln
305                 310                 315                 320

Gly Pro Ser Gly Asp Pro Gly Pro Pro Gly Pro Pro Gly Lys Glu Gly
                325                 330                 335

Leu Pro Gly Pro Gln Gly Pro Pro Gly Phe Gln Gly Leu Gln Gly Thr
            340                 345                 350

Val Gly Glu Pro Gly Val Pro Gly Pro Arg Gly Leu Pro Gly Leu Pro
        355                 360                 365

Gly Val Pro Gly Met Pro Gly Pro Lys Gly Pro Pro Gly Pro Pro Gly
    370                 375                 380

Pro Ser Gly Ala Val Val Pro Leu Ala Leu Gln Asn Glu Pro Thr Pro
385                 390                 395                 400

Ala Pro Glu Asp Asn Ser Cys Pro Pro His Trp Lys Asn Phe Thr Asp
                405                 410                 415

Lys Cys Tyr Tyr Phe Ser Val Glu Lys Glu Ile Phe Glu Asp Ala Lys
            420                 425                 430

Leu Phe Cys Glu Asp Lys Ser Ser His Leu Val Phe Ile Asn Thr Arg
        435                 440                 445

Glu Glu Gln Gln Trp Ile Lys Lys Gln Met Val Gly Arg Glu Ser His
    450                 455                 460

Trp Ile Gly Leu Thr Asp Ser Glu Arg Glu Asn Glu Trp Lys Trp Leu
465                 470                 475                 480

Asp Gly Thr Ser Pro Asp Tyr Lys Asn Trp Lys Ala Gly Gln Pro Asp
                485                 490                 495

Asn Trp Gly His Gly His Gly Pro Gly Glu Asp Cys Ala Gly Leu Ile
            500                 505                 510

Tyr Ala Gly Gln Trp Asn Asp Phe Gln Cys Glu Asp Val Asn Asn Phe
        515                 520                 525

Ile Cys Glu Lys Asp Arg Glu Thr Val Leu Ser Ser Ala Leu
    530                 535                 540
```

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 266 to nucleotide 1651;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 521 to nucleotide 1651;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 335 to nucleotide 634;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone as294_3 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone as294_3 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone as294_3 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone as294_3 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 123;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising eight consecutive amino acids of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone as294_3 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

8. The protein of claim 7, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

9. The protein of claim 7, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 123.

10. A composition comprising the protein of claim 7 and a pharmaceutically acceptable carrier.

11. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

12. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 262 to nucleotide 3096;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1118 to nucleotide 1527;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone aw92_1 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone aw92_1 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone aw92_1 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone aw92_1 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

13. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 287 to amino acid 422;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising eight consecutive amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone aw92_1 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

14. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

15. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 612 to nucleotide 806;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 744 to nucleotide 806;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 794;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bd316_2 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bd316_2 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bd316_2 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bd316_2 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

16. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 61;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising eight consecutive amino acids of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone bd316_2 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

17. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

18. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 7 to nucleotide 300;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 363;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk130_4 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk130_4 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bk130_4 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bk130_4 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

19. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising eight consecutive amino acids of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone bk130_4 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

20. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

21. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 52 to nucleotide 1863;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1219 to nucleotide 1863;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1099 to nucleotide 1743;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bv131_5 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bv131_5 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bv131_5 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bv131_5 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:10;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

22. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 430 to amino acid 564;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising eight consecutive amino acids of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone bv131_5 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

23. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

24. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 67 to nucleotide 690;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 1 to nucleotide 576;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bv227_1 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bv227_1 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bv227_1 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bv227_1 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:12;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

25. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 170;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising eight consecutive amino acids of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone bv227_1 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

26. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.

27. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 657 to nucleotide 1469;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 678 to nucleotide 1103;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cd265_11 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cd265_11 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cd265_11 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cd265_11 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:14;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

28. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 8 to amino acid 149;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising eight consecutive amino acids of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone cd265_11 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

29. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

30. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 261 to nucleotide 896;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 330 to nucleotide 896;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 515;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ej265_4 deposited under accession number ATCC 98444;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ej265_4 deposited under accession number ATCC 98444;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ej265_4 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ej265_4 deposited under accession number ATCC 98444;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:16;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

31. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 85;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising eight consecutive amino acids of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone ej265_4 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

32. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

33. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 946 to nucleotide 2232;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 1336 to nucleotide 1853;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ey29_8 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ey29_8 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ey29_8 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ey29_8 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:18;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

34. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 138 to amino acid 302;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising eight consecutive amino acids of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone ey29_8 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

35. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

36. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2588 to nucleotide 3439;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 3005 to nucleotide 3502;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gm114_10 deposited under accession number ATCC 98444;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gm114_10 deposited under accession number ATCC 98444;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gm114_10 deposited under accession number ATCC 98444;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gm114_10 deposited under accession number ATCC 98444;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:20;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

37. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 145 to amino acid 284;

(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising eight consecutive amino acids of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone gm114_10 deposited under accession number ATCC 98444;

the protein being substantially free from other mammalian proteins.

38. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

FIGURE 1A

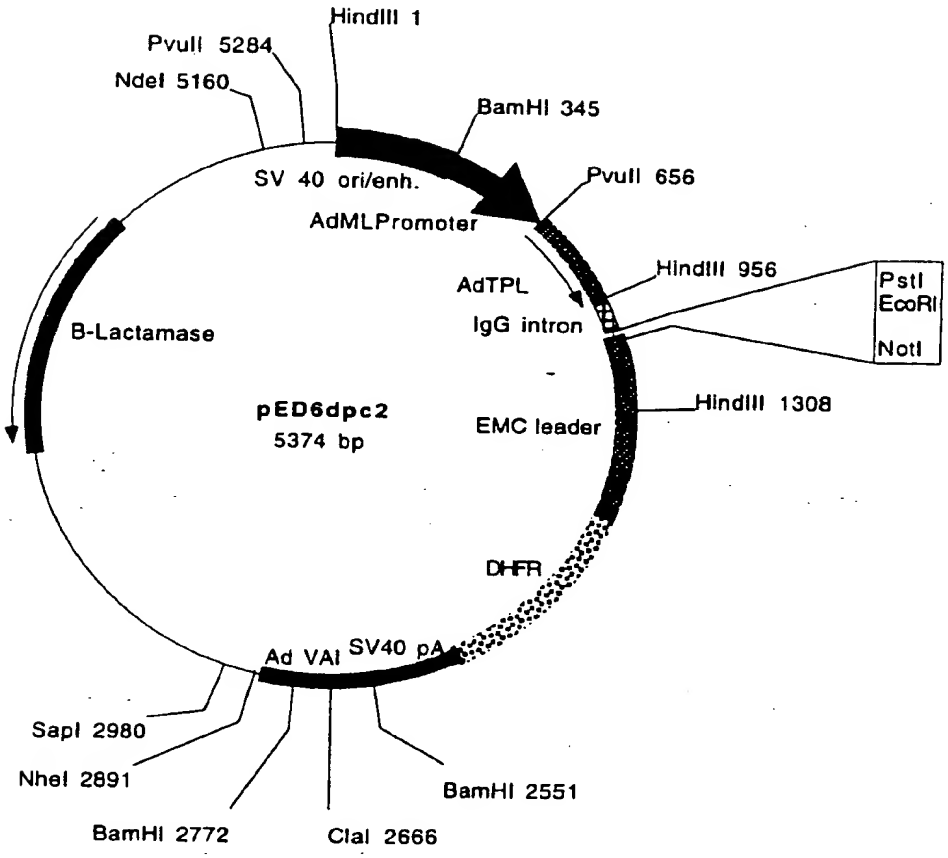
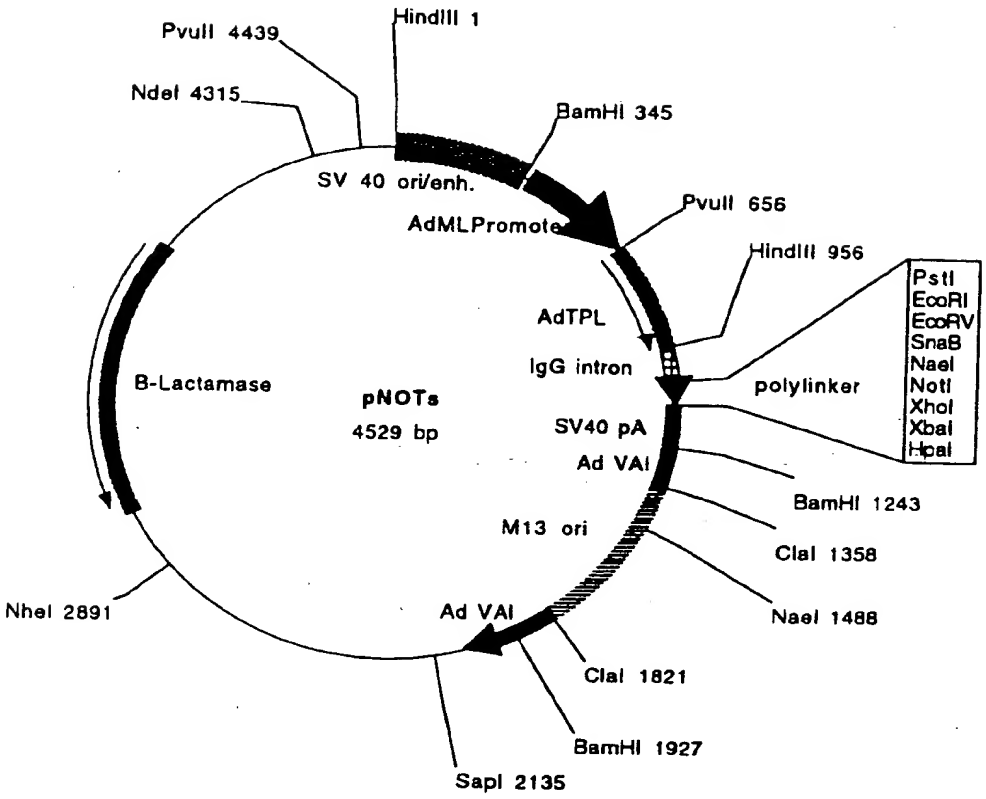


FIGURE 1B


PvuII 4439
HindIII 1
NdeI 4315
BamHI 345
SV 40 ori/enh.
PvuII 656
AdMLPromote
HindIII 956
PstI
EcoRI
EcoRV
SnaB
NaeI
NotI
XhoI
XbaI
HpaI
AdTPL
IgG intron
polylinker
B-Lactamase
pNOTs
4529 bp
SV40 pA
Ad VAI
BamHI 1243
M13 ori
ClaI 1358
NaeI 1488
NheI 2891
Ad VAI
ClaI 1821
BamHI 1927
SapI 2135

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, A61K 38/17

A3

(11) International Publication Number: **WO 98/55614**

(43) International Publication Date: 10 December 1998 (10.12.98)

(21) International Application Number: PCT/US98/11210

(22) International Filing Date: 1 June 1998 (01.06.98)

(30) Priority Data:

08/868,899	4 June 1997 (04.06.97)	US
08/868,898	4 June 1997 (04.06.97)	US
08/869,192	4 June 1997 (04.06.97)	US
08/869,191	4 June 1997 (04.06.97)	US
08/869,193	4 June 1997 (04.06.97)	US
08/868,697	4 June 1997 (04.06.97)	US
08/868,698	4 June 1997 (04.06.97)	US
08/868,900	4 June 1997 (04.06.97)	US
08/868,696	4 June 1997 (04.06.97)	US
08/869,194	4 June 1997 (04.06.97)	US
09/087,255	29 May 1998 (29.05.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Street, Newton, MA 02160 (US). McCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US). HOWES, Steven, H.; Apartment 2, 44 Chester Street, Somerville, MA 02144 (US). FECHTEL, Kim; 46 Marion Road, Arlington, MA 02174 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
1 April 1999 (01.04.99)

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/11210

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
3 E	WO 98 25959 A (CHIRON CORP. (US); ESCOBEDO - J; HU Q; GARCIA P; WILLIAMS L; KOTHAKOTA S) 18 June 1998 see page 3, line 21-24 see page 9, line 26-31 see page 10, line 24-29 see page 18, line 18-23 see page 19, line 30 - page 20, line 6 see page 20, line 28 - page 21, line 17 Seq.ID:15 see page 42 Seq.ID:34 see page 66 - page 68 see page 74 - page 77; claims --- -/--	1-7,10, 11

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 November 1998	15.02.1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/11210

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 98 45435 A (GENETICS INSTITUTE INC (US); JACOBS; MCCOY; LAVALLIE; RACIE ET AL.) 15 October 1998 see page 56, line 28 - page 57, line 6 Seq.ID:494 see page 259 ---	1-7,9-11
X	EP 0 679 716 A (MATSUBARA KENICHI; OKUBO KOUSAKU (JP)) 2 November 1995 cited in the application see page 8, line 20-35 see page 20, line 25-51 Seq.ID:2648 see page 907 - page 908 ---	1,11
X	Database EMBL Emest10, Entry HS673190 Accession number R88673 26 August 1995 97% identity with Seq.ID:1 nt.1-338 XP002084310 cited in the application see the whole document ---	1,11
X	Database EMBL Emest9, Entry HS37978 Accession number R15379 22 April 1995 97% identity with Seq.ID:1 nt.326-725 XP002084311 cited in the application see the whole document ---	1,11
X	Database EMBL Emest9, Entry HS273155 Accession number H18273 2 July 1995 98% identity with Seq.ID:1 nt.641-1065 XP002084312 see the whole document ---	1,11
X	Database EMBL Emest7, Entry HS1210943 Accession number AA405257 11 May 1997 99% identity with Seq.ID:1 nt.831-1339 XP002084313 see the whole document ---	1,11
X	Database EMBL Emest11, Entry HS796352 Accession number W56796 8 June 1996 98% identity with Seq.ID:1 nt.1319-1741 reverse orientation XP002084314 see the whole document --- -/--	1,11

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/11210

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 97 07198 A (GENETICS INSTITUTE INC (US); JACOBS K; MCCOY JM; KELLEHER K; CARLIN M) 27 February 1997 ---	
A	TASHIRO K. ET AL.: "Signal sequence trap: a cloning strategy for secreted proteins and type I membrane proteins" SCIENCE, vol. 261, 30 July 1993, pages 600-603, XP000673204 ---	
A	YOKOYAMA-KOBAYASHI M. ET AL.: "A signal sequence detection system using secreted protease activity as an indicator" GENE, vol. 163, 1995, pages 193-196, XP002053953 ---	
A	US 5 536 637 A (JACOBS KENNETH; GENETICS INSTITUTE INC. (US)) 16 July 1996 cited in the application -----	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 98/11210

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-11

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-11

Polynucleotide comprising the nucleotide sequence of Seq.ID:1 and encoding a polypeptide of Seq.ID:2 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Host cell transformed with said polynucleotide. Protein comprising an amino acid sequence of Seq.ID:2 or fragments thereof. Process for producing said protein.

2. Claims: 12-14

Polynucleotide comprising the nucleotide sequence of Seq.ID:3 and encoding a polypeptide of Seq.ID:4 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Protein comprising an amino acid sequence of Seq.ID:4 or fragments thereof.

3. Claims: 15-17

As invention 2 but concerning Seq.ID:5 and 6.

4. Claims: 18-20

As invention 2 but concerning Seq.ID:7 and 8.

5. Claims: 21-23

As invention 2 but concerning Seq.ID:9 and 10.

6. Claims: 24-26

As invention 2 but concerning Seq.ID:11 and 12.

7. Claims: 27-29

As invention 2 but concerning Seq.ID:13 and 14.

8. Claims: 30-32

As invention 2 but concerning Seq.ID:15 and 16.

9. Claims: 33-35

As invention 2 but concerning Seq.ID:17 and 18.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: 36-38

As invention 2 but concerning Seq.ID:19 and 20.

INTERNATIONAL SEARCH REPORT

...mation on patent family members

Intern...nal Application No
PCT/US 98/11210

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9825959 A	18-06-1998	AU 5796298 A	03-07-1998
WO 9845435 A	15-10-1998	AU 6956698 A	30-10-1998
EP 0679716 A	02-11-1995	AU 8116494 A	13-06-1995
		CA 2153480 A	01-06-1995
		WO 9514772 A	01-06-1995
WO 9707198 A	27-02-1997	US 5707829 A	13-01-1998
		AU 6712396 A	18-02-1997
		AU 6768596 A	12-03-1997
		CA 2227220 A	06-02-1997
		CA 2229208 A	27-02-1997
		EP 0839196 A	06-05-1998
		EP 0851875 A	08-07-1998
		WO 9704097 A	06-02-1997
US 5536637 A	16-07-1996	US 5712116 A	27-01-1998